Technische Universität München

Ingenieurfakultät Bau Geo Umwelt

Lehrstuhl für Siedlungswasserwirtschaft

Development of regression models to predict biogas production rate and biogas yield

Vasileios Dandikas

Vollständiger Abdruck der von der Ingenieurfakultät Bau Geo Umwelt der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzende:	apl. Prof. Dr. rer. nat. habil. Brigitte Helmreich
Prüfer der Dissertation:	1. Prof. Dr.-Ing. Jörg E. Drewes
	2. Assoc. Prof. Dr.-Ing. Jin Mi Triolo
	3. Prof. Dr. agr. Hauke Heuwinkel

Die Dissertation wurde am 26.06.2018 bei der Technischen Universität München eingereicht und durch die Ingenieurfakultät Bau Geo Umwelt am 30.10.2018 angenommen.

Preface and acknowledgements

This dissertation consists of four already published articles in peer reviewed journals. All four articles present results of research on anaerobic digestion process and more specific on biomethane potential (BMP) tests.

The biogas yield and the biogas production rate are important parameters for assessing any biodegradable material in order to define its suitability as feedstock in a biogas plant. The aims of this study were to identify and quantify the mathematical relationship between the chemical compounds (fodder analysis) and BMP of feedstocks, and to develop a model for predicting the biogas and methane yield. Moreover, this thesis notes the possibilities and limitations of linear regression models for BMP prediction and proposed an approach for biogas yield and biogas production rate prediction suitable for researches and practitioners.

At this point, I would like to express my deeply gratitude to all who supported me and helped me to accomplish this thesis.

First of all, I am deeply thankful to my supervisor Prof. Dr.-Ing. Jörg E. Drewes for his trust on me and his valuable guidance. Furthermore, I would like to express my fully gratitude to Prof. Dr. agr. Hauke Heuwinkel for his willingness to discuss all the arisen issues during my research and offering suggestions for improvement. I also wish to thank Assoc. Prof. Dr.-Ing. Jin Mi Triolo for her willingness to participate at my Ph.D. defense and for her advice. It was a privilege and an honor to work with them; they help me to extent my scientific knowledge.

I am very grateful to Dr.-Ing. Konrad Koch for his advice, support and constructive suggestions, which were very important for the accomplishment of this study. Furthermore, I appreciate our friendship developed throughout these years.

I am thankful to Dr. Georg Wendl, Dr. Stefan Neser und Dr. Fabian Lichti for their support and encouragement at the Bavarian State Research Center for Agriculture (LfL) in Freising, Germany. I would also like to thank Dr. Stephan Hartmann, Dorothea Hofmann, Dr. Joachim Eder, Monika Fleschhut, Dr. Maendy Fritz und Dr. Anja Hartmann for their support.

I am very thankful to my “batch-team” (Anke Aschmann, Natascha Siddiqui and Michael Hanrieder) for their excellent assistance and to my colleagues at Institute for Agricultural Engineering and Animal Husbandry (ILT) in Freising for the wonderful working environment, their suggestions and encouragement.

Moreover, I am totally grateful to my family for their encouragement and deeply support. Their love gives me strength and motivates me.

This study was funded by the Bavarian State Ministry of Food, Agriculture and Forestry and the Bavarian State Ministry of Economic Affairs and Media, Energy and Technology. Project codes: EW/12/04 and BE/15/06.

Abstract

In recent years, the biogas sector has grown on a global level as a result of the sustainable and low emission energy policy. Anaerobic digestion is considered a cornerstone of both energy transition and circular economy. Moreover, biogas is advantageous compared to other energy sources due to the fact that it can be stored and utilized independent of time and place at times of higher energy demand. Two concepts for demand-driven flexible operations of a biogas plant have been suggested, based on either the storage of biogas or on-demand biogas production by means of the feedstocks.

Energy crops (first generation biomass) are currently the main feedstock of agricultural biogas plants. However, due to increasing land use for the production of energy crops, they have been critically discussed (fuel vs. food debate), and the public acceptance of the agricultural biogas plants has been negatively affected. Thus, the utilization of second generation biomass (non-food material) is desirable. To determine their suitability as feedstocks of a biogas plant many factors should be taken into account; one of them is the biomethane potential (BMP).

The BMP and the degradation rate of feedstocks can be determined by standardized anaerobic digestion batch tests (BMP tests). However, those tests are very complex, costly, time-consuming and their standardization is still very challenging. The experimental procedure of BMP tests is well described in different guidelines and methods books, but round robin tests have revealed a high deviation of measured methane yield from the same substrate among different laboratories.

Thus, the aims of this study were: a) to identify and quantify the mathematical relationship between the chemical compounds (fodder analysis) and BMP of feedstocks and b) to develop a model for predicting the biogas and methane yield. In addition, it attempted to describe the complex biological process based on the chemical composition of plants and a first-order reaction.

In this study, fodder analysis and BMP tests with high temporal resolution were performed in order to identify statistical correlations between the biogas production rate, biogas yield and the chemical composition of various energy crops. Different species and cultivars of energy crops were analyzed in order to develop a broadly applicable regression model. The crops were collected at varying developmental stages between 2012 and 2016. Based on those samples, four datasets were developed in order to statistically investigate the effect of the plant's chemical compounds on the biogas production.

The results showed that the acid detergent lignin (ADL) content of various energy crops had a strong negative correlation with biogas and methane yields, and for a mono-causal

regression model, the ADL was the only candidate of the chemical compounds. Based on regression analysis, more than 80 % of the samples' variation in BMP could be explained by ADL. Apart from ADL, the principal component analysis showed that hemicellulose (HC) was statistically important for biogas yield prediction. As no other variables yielded significant influence, ADL and HC are suggested as the suitable variables to predict biogas yields across different species of energy crops, with a prediction error for the calibration of about 8 %. However, the ADL content of a specific plant group (grassland samples) dataset could not explain the samples' variation in BMP. The estimation error of the global model for the prediction of biogas yield from grassland plant species was about 12 %. For high accuracy in predicting the biogas yield (prediction error 5 % for the calibration), a third regressor was needed and crude protein (XP) was added. The results showed that a global model can predict the variation of BMPs among the different plant species, but in order to precisely predict the variation on BMP among the plant cultivars, a specific (local) model is needed.

Besides biogas yield prediction, the kinetics of biogas production during a BMP test was described. This was accomplished by investigating the statistical correlations between the hydrolysis rate constant (k_h) and the chemical composition of various energy crops. Results indicated that the analytical parameters non-fiber carbohydrates (NFC) and XP were statistically suitable for the prediction of k_h. The regressors of k_h prediction model and the regressors of BMP prediction model are different; this observation indicates that the biogas production rate is not necessary correlated with the biogas yield. Furthermore, a first-order kinetic model and the proposed regression models can be utilized for the prediction of the biogas yield and biogas rate in a BMP test.

Finally, an independent dataset was used to assess and validate previously published prediction models and those developed in this study, as well as to expose the power and limitation of linear regression models for BMP prediction. The results revealed that linear regression models suitably depict the variation in the biogas yield to get a substrate ranking. However, the prediction error for the absolute values may be high since systematic external effects cannot be identified from the regression models. Despite the fact that the prediction is limited due to its simplicity and cannot accommodate extreme cases, this approach can be a useful tool for practitioners in order to assess different feedstocks for biogas production. The study proposes a novel approach for the prediction of both biogas yield and biogas production rate simultaneously based only on fodder analysis.

Zusammenfassung

In den letzten Jahren ist die Biogasbranche durch die nachhaltige und emissionsarme Energiepolitik weltweit gewachsen. Die anaerobe Vergärung hat sich als Eckpfeiler der Energiewende und der Kreislaufwirtschaft etabliert. Darüber hinaus bietet Biogas im Vergleich zu anderen Energieträgern den entscheidenden Vorteil, dass es zeit- und ortsunabhängig gespeichert und zu Zeiten höheren Energiebedarfs abgerufen werden kann. Es wurden zwei Konzepte für die bedarfsorientierte, flexible Stromerzeugung einer Biogasanlage vorgeschlagen, die Biogasspeicherung oder die variable Biogaserzeugung durch die Einsatzstoffe.

Energiepflanzen (Biomasse der ersten Generation) sind derzeit der wichtigste Einsatzstoff für landwirtschaftliche Biogasanlagen. Aufgrund der zunehmenden Landnutzung für die Produktion von Energiepflanzen wurde deren Einsatz in der Öffentlichkeit jedoch lange kontrovers diskutiert (Teller-Tank-Diskussion) und führte zur aktuell geringen Akzeptanz landwirtschaftlicher Biogasanlagen in der Bevölkerung. Folglich gewinnt die Nutzung von Biomasse der zweiten Generation (Rest- und Abfallstoffe) immer mehr an Bedeutung. Um ihre Eignung als Einsatzstoffe einer Biogasanlage zu ermitteln, müssen viele Faktoren berücksichtigt werden, einer davon ist das Biomethanpotenzial (BMP).

Das BMP und die Abbaurate der Einsatzstoffe können durch anaerobe Batchversuche (BMP-Tests) bestimmt werden. Diese Untersuchungen sind jedoch sehr komplex, kostenintensiv, zeitaufwändig und ihre Standardisierung ist immer noch sehr anspruchsvoll. Das experimentelle Verfahren der BMP-Tests ist in verschiedenen Richtlinien und Methodenbüchern gut beschrieben, allerdings haben Ringversuche eine hohe Abweichung in der gemessenen Methanausbeute desselben Substrates zwischen verschiedenen Laboren ergeben.

Daher waren die Ziele dieser Studie: a) den mathematischen Zusammenhang zwischen den Inhaltsstoffen (Futtermittelanalytik) und dem BMP von Einsatzstoffen zu identifizieren und zu quantifizieren, und b) ein Modell zur Vorhersage der Biogas- und Methanausbeute zu entwickeln. Darüber hinaus wurde versucht, den komplexen biologischen Prozess basierend auf der chemischen Zusammensetzung der Pflanzen und einer Reaktion erster Ordnung zu beschreiben.

In dieser Studie wurden Futtermittelanalysen und BMP-Tests mit hoher zeitlicher Auflösung durchgeführt, um statistische Korrelationen zwischen Biogasproduktionsrate, Biogasausbeute und der chemischen Zusammensetzung verschiedener Energiepflanzen zu identifizieren. Verschiedene Arten und Sorten von Energiepflanzen wurden analysiert, um ein breit anwendbares Regressionsmodell zu entwickeln. Die Pflanzen wurden zwischen 2012 und 2016 in unterschiedlichen Entwicklungsstadien geerntet. Anhand dieser Proben wurden vier Datensätze entwickelt, um die Wirkung der

chemischen Verbindungen der Pflanzen auf die Biogasproduktion statistisch zu untersuchen.

Die Ergebnisse zeigten, dass der Gehalt an Säure-Detergenz-Lignin (ADL) verschiedener Energiepflanzen eine starke negative Korrelation mit den Biogas- und Methanausbeuten aufweist und ADL für ein monokausales Regressionsmodell der einzige Kandidat der chemischen Verbindungen war. Basierend auf einer Regressionsanalyse konnten mehr als 80 % der Probenvariation an BMP durch ADL erklärt werden. Eine Hauptkomponentenanalyse zeigte, dass neben ADL nur Hemizellulose (HC) für die Vorhersage der Biogasausbeute statistisch geeignet war. Da keine anderen Variablen einen signifikanten Einfluss auf die Biogasausbeute hatten, wurden ADL und HC als geeignete Variablen zur Vorhersage der Biogasausbeute von unterschiedlichen Pflanzenarten vorgeschlagen. Der Vorhersagefehler der Kalibrierung lag bei 8 %. Der ADL-Gehalt eines spezifischen Pflanzengruppen-Datensatzes (Grünlandproben) konnte jedoch die Variation der Proben im BMP nicht erklären. Für die Grünlandproben lag der Vorhersagefehler des globalen Modells bei 12 %. Für eine höhere Vorhersagegenauigkeit der Biogasausbeute (Vorhersagefehler 5 % bei der Kalibrierung) wurde ein dritter Regressor nämlich das Rohprotein (XP) hinzugefügt. Die Ergebnisse zeigten, dass ein globales Modell die Variation der BMPs zwischen den verschiedenen Pflanzenarten vorhersagen kann. Um allerdings die Variation der BMPs der Pflanzensorten präzise vorhersagen zu können, wird ein spezifisches (lokales) Modell benötigt.

Neben der Vorhersage der Biogasausbeute wurde die Kinetik der Biogasproduktion während eines BMP-Tests beschrieben. Dazu wurden die statistischen Zusammenhänge zwischen der Hydrolysekonstante (k_h) und der chemischen Zusammensetzung verschiedener Energiepflanzen untersucht. Die Ergebnisse zeigten, dass die analytischen Parameter Nicht-Faser-Kohlenhydrate (NFC) und XP für die Vorhersage von k_h statistisch geeignet waren. Die Regressoren des k_h-Vorhersagemodells und die Regressoren des BMP-Vorhersagemodells sind unterschiedlich. Diese Beobachtung zeigt, dass die Biogasrate nicht zwingend mit der Biogasausbeute korreliert. Darüber hinaus können ein kinetisches Modell erster Ordnung und die vorgeschlagenen Regressionsmodelle zur Vorhersage der Biogasausbeute und der Biogasrate in einem BMP-Test verwendet werden.

Schließlich wurde ein unabhängiger Datensatz verwendet, um bereits veröffentlichte und die in dieser Studie entwickelten Vorhersagemodelle zu bewerten und zu validieren. Darüber hinaus wurden das Potenzial und die Grenzen linearer Regressionsmodelle zur BMP-Vorhersage aufgezeigt. Die Ergebnisse zeigten, dass lineare Regressionsmodelle die Variation der Biogasausbeute für ein Substrat-Ranking geeignet abbilden. Der Vorhersagefehler für die absoluten Werte kann jedoch hoch sein, da systematische externe Effekte nicht aus den Regressionsmodellen identifiziert werden können. Trotz der Tatsache, dass die Vorhersage aufgrund ihrer Einfachheit begrenzt ist und extreme

Fälle nicht berücksichtigen kann, kann dieser Ansatz ein nützliches Werkzeug für die Praxis sein, um verschiedene Einsatzstoffe für die Biogasproduktion zu bewerten. Diese Studie schlägt ein neues Konzept zur gleichzeitigen Vorhersage der Biogasausbeute und der Biogasproduktionsraterate auf Basis der Futtermittelanalyse vor.

Published research articles

This dissertation is based on already published papers in peer-reviewed scientific journals. The following papers are included in the appendix of this thesis.

Paper I

Dandikas, V., Heuwinkel, H., Lichti, F., Drewes, J.E., Koch, K., (2014). Correlation between biogas yield and chemical composition of energy crops. Bioresource Technology 174, 316–320. doi:10.1016/j.biortech.2014.10.019

Paper II

Dandikas, V., Heuwinkel, H., Lichti, F., Drewes, J.E., Koch, K., (2015). Correlation between Biogas Yield and Chemical Composition of Grassland Plant Species. Energy & Fuels 29, 7221–7229. doi:10.1021/acs.energyfuels.5b01257

Paper III

Dandikas, V., Heuwinkel, H., Lichti, F., Eckl, T., Drewes, J.E., Koch, K., (2018). Correlation between hydrolysis rate constant and chemical composition of energy crops. Renewable Energy 118, 34–42. doi.:10.1016/j.renene.2017.10.100

Paper IV

Dandikas, V., Heuwinkel, H., Lichti, F., Drewes, J.E., Koch, K., (2018). Predicting methane yield by linear regression models: A validation study for grassland biomass. Bioresource Technology 265, 372–379. doi.:10.1016/j.biortech.2018.06.030

Topic related scientific contributions

Publications:

Dollhofer V.; **Dandikas V**.; Dorn-In S.; Bauer C.; Lebuhn M.; Bauer J. (2018) Accelerated biogas production from lignocellulosic biomass after pre-treatment with Neocallimastix frontalis. Bioresource Technology 264, 219–227.

Rath J., Herrmann A.; Heuwinkel H.; **Dandikas V.**; Lichti F. (2017). Welcher Maissortentyp für die Biogasanlage? Biogas Journal, 01-2017, 78 – 84, pp 7, ISSN 1619-8913.

Rath J., Herrmann A.; Heuwinkel H.; **Dandikas V.**; Lichti F. (2016). Welcher Maissortentyp für die Biogasanlage? - Wechselspiel der Inhaltsstoffe beeinflusst das Potenzial der Biogasausbeute. Mais, 04_2016, 171 - 175, pp 5, ISSN 0341-5155.

Dollhofer V., Nast M., Kinker I., Dorn-In S., **Dandikas V.**, Bauer J. and Lebuhn M. (2015). Anaerobe Pilze im Biogasprozess (Eng.: Anaerobic fungi in the biogas process). In proceedings of the 4. KTBL/FNR- Kongress „Biogas in der Landwirtschaft – Stand und Perspektiven“, Potsdam Germany, September 22 - 23 2015. pp 11.

Dandikas V., Heuwinkel H., Lichti F., Drewes J. E., and Koch K. (2015); Development of an empirical model to estimate the biogas yield of energy crops. In proceedings of HEZagrar PhD Symposium, Freising, Germany, April, 2015. pp 2.

Dandikas V., Heuwinkel H., Lichti F., Drewes J. E., and Koch K. (2015); Prediction of biogas yield based on the chemical composition: Potential and limitations. In proceedings of Conference on Monitoring & process control of anaerobic digestion plants. Leipzig, Germany, March 17-18, 2015, 2014. pp 1.

Dandikas V., Heuwinkel H., Lichti F., Drewes J. E., and Koch K. (2014); Influence of chemical composition on potential biogas yield of lignocellulosic biomass. In proceedings of Biogas Science 2014, International Conference on Anaerobic Digestion. Vienna, Austria, October 26 – 30, 2014. pp 1.

Weber, A., Dahlhoff, A., **Dandikas, V.**, Effenberger, M., Neser, S. (2013); 2.4 Energetische Verwertung, in: Handbuch Mais: Grundlagen • Anbau • Verwertung • Ökonomie. DLG-Verlag, Frankfurt am Main, ISBN 978-3-7690-0826-5.

Presentations:

Dandikas, V.: 'Berechnung der Biogasausbeute und der Hydrolysekonstante', Freising, 29.06.2017, Biogas Forum, Bayerische Landesanstalt für Landwirtschaft.

Dandikas, V.: 'Entwicklung eines mathematischen Modells zur Abschätzung des Biogasertragspotentials von NawaRo'. München, 09.12.2015, Biogas Jour Fixe, Bayerisches Staatsministerium für Ernährung, Landwirtschaft und Forsten.

Dandikas, V.: 'Development of an empirical model to estimate the biogas yield of energy crops ', Freising, 21.04.2015, 1. HEZagrar PhD Symposium, Technische Universität München.

Dandikas, V.: 'Prediction of biogas yield based on the chemical composition: Potential and limitations', Leipzig, 17.03.2015, Deutsches Biomasseforschungszentrum.

Dandikas, V.: 'Batchformel – Modellentwicklung anhand bestehender und neuer Datensätze', München, 09.12.2014, Biogas Jour Fixe, Bayerisches Staatsministerium für Ernährung, Landwirtschaft und Forsten.

Dandikas, V.: 'Influence of chemical composition on potential biogas yield of lignocellulosic biomass', Wien, Österreich, 29.10.2014, Biogas Science 2014, Universität für Bodenkultur Wien.

Posters:

Hartung, C.; Heuwinkel, H.; **Dandikas, V.**: Biogasertragspotential von Paludikultur-Pflanzen. 25.-26.09.2017, Klimaschutz und Moornutzung: Potentiale in Deutschland, Ernst-Moritz-Arndt Universität Greifswald.

Dandikas, V., Heuwinkel, H.; Lichti, F.; Drewes, J. E.; Koch, K.: 'Comparing a global and a local modeling approach for the prediction of the biogas yield of energy crops', Hohenheim, 08.03.2017, Progress in Biogas IV, Universität Hohenheim.

Dandikas, V.: Biomethanpotenzial zur Substratbewertung und –charakterisierung, LfL-Infostand.

Abbreviations

ADF	Acid detergent fiber
ADL	Acid detergent lignin
AQU	Abteilung Qualitätssicherung und Untersuchungswesen (Central Department for Quality Assurance and Analytics)
BBCH	Biologische Bundesanstalt, Bundessortenamt und Chemische Industrie (Code of the phenological growth stages of plants)
BMP	Biomethane potential
CH_4	Methane
CL	Cellulose
CO_2	Carbon dioxide
CVRMSE	Coefficient of variation of root-mean-square error
EEG	Erneuerbaren-Energien-Gesetz (Renewable Energy Act)
FM	Fresh matter
FOM	Fermentable organic matter
GHG	Greenhouse gas
H_2	Hydrogen gas
H_2S	Hydrogen sulfide
HC	Hemicellulose
ILT	Institut für Landtechnik und Tierhaltung (Institute for Agricultural Engineering and Animal Husbandry)
IPZ	Institut für Pflanzenbau und Pflanzenzüchtung (Institute for Crop Science and Plant Breeding)
LfL	Bayerische Landesanstalt für Landwirtschaft (Bavarian State Research Center for Agriculture)
MLR	Multiple linear regression
N	Nitrogen
n	Sample size
NDF	Neutral detergent fiber
NFC	Non-fiber carbohydrates
NfE	Nitrogen free extract
NH_4^+-N	Ammonium nitrogen
NIRS	Near-infrared spectroscopy
OR	Organic residue
p	Probability
PC	Principal component
PCA	Principal component analysis
PCR	Principal component regression
r	Correlation coefficient

R^2	Coefficient of determination
RMSE	Root-mean-square error
ST	Starch
TMR	Total mixed ration
TS	Total solids
VDI	Verein Deutscher Ingenieure (Association of German Engineers)
VDLUFA	Verband Deutscher Landwirtschaftlicher Untersuchungs- und Forschungsanstalten (Association of German Agricultural Analytic and Research Institutes)
VQ	Quotient of digestibility
VS	Volatile solids
XA	Crude ash
XF	Crude fibers
XL	Crude lipids
XP	Crude protein
XS	Reduced sugar
Y_B	Biogas yield
Y_M	Methane yield

Table of contents

1 Introduction

In order to reduce the anthropogenic contribution to greenhouse gas (GHG) emissions and global warming, sustainable energy policy and low emission energy production are needed. In 2010, the European Commission proposed a ten-year strategy and set targets in order to accelerate sustainable growth (Bere et al., 2015). For each member state, specific targets are defined concerning the reduction of GHG emissions, the increase of the share of renewable energy and the increase of energy efficiency, which must be achieved by the year 2020 (Zervos et al., 2011). In Germany, for instance, the share of renewable energy should reach at least 18 % of the total energy consumption. Moreover, institutions from across the political and scientific spectrum state the importance of acting against the environmental damage and new targets were set in the 2015 Paris climate agreement (Liobikienė and Butkus, 2017).

Electricity production from fossil fuels is the main source of GHG emissions. By switching to renewable energy sources, the impact on GHGs can be reduced or, in some cases eliminated. The utilization of biogas allows for less dependence on fossil fuels and will reduce GHG emissions (Meyer-Aurich et al., 2012; Uusitalo et al., 2013). However, due to the increasing allocation of agricultural land for energy crop production, the use of energy crops has been critically discussed. For instance, the increasing maize cultivation in Germany has damaged the public acceptance of biogas plants (Kortsch et al., 2015). Recent studies have shown that there is a sufficient quantity of sustainable alternatives to the use of maize in all the member states of the EU-28 to ensure continuous development of the European biogas sector (Meyer et al., 2018). These alternatives include utilizing biomass which cannot be used as feed or food, so called second generation biomass. A desirable example of this is non-agricultural grasslands and roadside grass as they exhibit a high potential of biomass.

Methane production from second generation biomass leads to the sustainable energy generation avoiding the “food versus fuel” debate (Chandra et al., 2012). The produced biogas can have a variety of potential uses, e.g. as part of a co-generation energy system for electricity and heat production. Moreover, it can be utilized to fill electricity production gaps (Ellabban et al., 2014; Esen and Yuksel, 2013).

The number of biogas plants in Germany started to rise in 2004 with the political support of the Renewable Energy Act (EEG). As a result, the production of renewable raw materials or energy crops (especially maize silage) as feedstock for biogas plants has also increased. Biogas plants have a high impact on regional agriculture in Germany, as maize is currently the most important energy crop for biogas production at agricultural biogas plants. However, the renewable energy policy in 2012 introduced an amendment to the EEG act. Among other strategies, a maximum amount of 60 % on the

share of maize and cereal grain kernels was introduced to prevent monoculture. Therefore, the need for alternative feedstocks is increasing.

In Germany there is regional grassland available that is not needed for livestock farming. Furthermore, agricultural residues (e.g. maize stover) and perennial plant species (e.g. cup plant) can also be utilized. These types of biomass could be used for biogas production, as they exhibit a significant energetic potential.

The determination of the biogas and methane yield of energy crops is a central element in biogas technology and a key parameter for substrate assessment and biogas plant planning and design. The biomethane potential (BMP) can be determined with lab-scale experiments. However, the practical use of these experiments is limited by the costly nature of BMP tests (ca. 200 - 500 € per sample) due to the expensive technical equipment, labor effort and long duration of the tests (ca. 25 - 45 d). Moreover, the BMP tests are biological systems and their standardization is very challenging. Due to this, it is difficult to ensure repeatability and reproducibility of the data among different laboratories (Raposo et al., 2011). Although a number of standard protocols are available (Angelidaki et al., 2009; Holliger et al., 2016; VDI, 2016; VDLUFA, 2011), the experimental setup is not always comparable.

A modeling approach based on the fodder analysis could provide quick and reliable information about the biogas and methane yield potential of different substrates. Fodder analyses are well-established standard methods based on analytical chemistry with high repeatability and reproducibility. The complexity is further reduced with the use of near-infrared spectroscopy (NIRS).

Several authors have developed a mathematical estimation for BMP prediction. However the effect of each compound on the BMP differs. For instance, crude lipids (XL) have been reported to have both a positive and a negative effect on biogas yield (Amon, 2007; Amon et al., 2007a). Crude fibers (XF) have also been reported to have sometimes a positive and sometimes a negative effect on biogas yield (Weißbach, 2008). Kaiser (2007) reported a positive effect of cellulose (CL) on BMP, while Triolo et al. (2011) reported a negative effect.

In order to expose the effect of the chemical compounds of the plant on the biogas yield and biogas production rate, defined plant species and varieties were collected at different developmental stages. During the field experiments, changes in composition of the chemical compounds were systematically recorded. A wide composition range of the chemical compounds were studied to expose the mono-causal effect and the interaction effect of the compounds on the biogas yield. In total, 312 samples were collected and tested in batch fermentation trials (BMP tests) under defined laboratory conditions; their chemical composition were analyzed (i.e. fodder analysis) as well. Based on statistical tools, specific chemical compounds were selected as regressors in

order to develop mathematical models that can predict the biogas production and degree of degradation of a feedstock.

The aim of this study was to develop a mathematical model which can be utilized to estimate the BMP test results quickly and cost-effectively. The study can be divided into four parts. The first part aimed to develop an across plant-species (global) model to identify the key variables for the prediction of biogas yield. The second part of the study aimed to minimize the estimation error by a plant group specific (local) model. In the third part, the effect of the chemical composition of energy crops on the hydrolysis rate constant was studied and a prediction model was developed. Finally, in fourth part BMP prediction models (two models from the literature and the developed ones) were assessed and validated to prove the accuracy potential of each modeling approach (global and local) and to compare with existing ones.

2 State of the knowledge

2.1 Biogas production process

Anaerobic digestion (AD) is a complex microbiological process that takes place in four steps (Fig. 1). Although the four groups of microorganisms have a different optimum growth range (Bauer et al., 2009), the biogas process is only stable when all four steps are in equilibrium due to the synergetic effects of the microbes. The process involves a broad variety of microorganisms, bacteria and archaea. The population dynamics of the microorganisms are dependent mainly on the environmental conditions in biogas reactors and on the feedstocks used (Blasco et al., 2014; Karakashev et al., 2005; Lebuhn et al., 2014). However, the population dynamics can be changed by lag of trace elements or other process inhibitory factors (Munk et al., 2010; Pobeheim et al., 2011).

The AD process is divided into four steps: hydrolysis, acidogenesis, acetogenesis and methanogenesis. In hydrolysis, complex organic substances are broken down into their monomers by hydrolytic bacteria or anaerobic fungi (Dollhofer et al., 2017). During this enzymatic process, particulate organic material is decomposed to solutes capable of being actively or passively transported across cell membranes before they can be microbially metabolized. Insoluble complex molecules (carbohydrates, proteins and lipids) are degraded into sugars, fatty acids and amino acids. Then, the acidogenesis takes place through acid-forming bacteria, producing mainly volatile fatty acids and alcohols. In the next step, acetogenesis, the anaerobic oxidation of long-chain fatty acids takes place to produce mainly acetic acid and hydrogen. The last step, methanogenesis, is the formation of methane by the hydrogenotrophic and acetoclastic methanogens.

Biogas formation is a process in which microbial and chemical aspects are closely linked (Angelidaki et al., 2009; Mittweg et al., 2012). Moreover, the chemical composition, in particular the lignocellulosic matrix, of the different plant species can affect the degree of degradation and the BMP (Li et al., 2013; Lübken et al., 2010). Under uninhibited anaerobic digestion conditions, hydrolysis of complex organic particulate material can be considered as the rate-limiting step of the anaerobic process (Pavlostathis and Giraldo-Gomez, 1991).

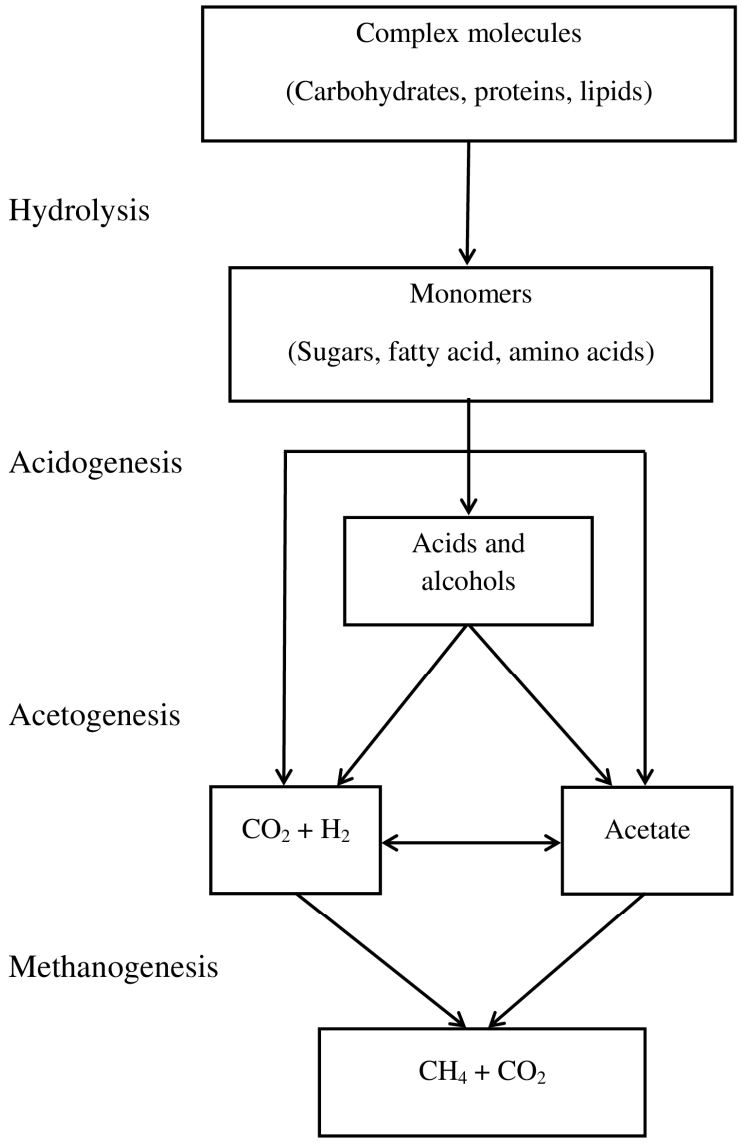

Figure 1: Flow diagram of the biogas process (Gerardi, 2003).

The AD process efficiency depends on substrate characteristics and operational conditions. The anaerobic digestion model no. 1 (ADM1) simulates the whole AD process including multiple steps of biochemical, as well as physico-chemical processes. Although modeling of the biogas production process is very challenging, it allows a better understanding and visualization of the biochemical processes and can optimize the performance of biogas plants.

2.2 Biogas and methane yields

The biogas yield describes the ultimate biogas production per amount of substrate added under defined conditions. The methane content in biogas will be analyzed to define the methane yield. Although, for energy production only the methane amount can be utilized, the biogas amount has to be known in order to undertake biogas plant design and operation. To achieve a more efficient operation of biogas plants, it is important to determine the biogas yield in advance. Anaerobic batch test (BMP test) under laboratory conditions make this possible, but a strict standardization is needed to assure inter- and intra-laboratory reproducibility (Angelidaki et al., 2009; Koch and Drewes, 2014).

BMP tests allow also quantifying the actual microbiological activity indirectly, according to the hourly biogas production. Since the BMP tests are performed under defined laboratory conditions, a comparison of the hourly production among the samples is possible. Hydrolysis is the rate limiting step of the whole anaerobic process and hydrolysis rate constant can be used as a variable for quantifying the velocity of the degradation of feedstock.

2.3 Experimental determination of biogas and methane yields

The biogas and methane yields of a substrate can be determined by discontinuous co-digestion under defined laboratory conditions. The Association of German Engineers (VDI) has developed guideline VDI 4630 for the determination of biogas and methane yields of organic substances (VDI, 2006). The main process conditions for a standardized batch test according to the guideline VDI 4630 are listed as follows: The particle size of the sample should be less than 10 mm; the volatile solids (VS) of the inoculum should be over 50 % of the total solids (TS); the inoculum should be degassed at the test temperature for a week; 1.5 to 2 % of the digester's working volume should be VS from the inoculum; the amount of VS from the substrate should be less than half the amount of the VS from the inoculum; microcrystalline cellulose should be used as a positive control, since it is 100 % degradable in a batch test; and the experiment should be terminated only when the daily rate of biogas falls below 1 % of the biogas production produced by then.

Angelidaki et al. (2009) proposed a protocol for the determination of the methane potential of solid organic wastes and energy crops in BMP tests. The authors highlighted the importance of the particle size of the test sample, the quality of inoculum, and the experimental setup.

In 2011, the Association of German Agricultural Analytic and Research Institutes (VDLUFA) published a method for the determination of biogas and methane yields of

agricultural biomass (VDLUFA, 2011). In the VDLUFA method some process conditions were defined more specifically with respect to the batch test of agricultural substrates. The stricter conditions are presented as follows: The inoculum should be biologically active material from a biogas plant and should have a total acid concentration below 500 mg acetic acid-equivalent per liter; the weighing accuracy should be at least 1 %; the TS content should be less than 10 % in the batch digester; the test temperature is defined at 37 ± 2 °C; and the biogas produced from inoculum alone should not be more than 20 % of the total biogas production of the sample with the inoculum. Moreover, to decide if the experiment ran properly, the absolute difference on biogas yield among the replicates of the test sample should be considered in addition to the biogas yield (validity of the measurement).

Although several norms and guidelines for BMP tests exist, inter-laboratory tests regularly still show a high variability of biogas and methane yields for the same substrate. To address this, Holliger et al. (2016) proposed that not only inter-laboratory but also intra-laboratory tests are needed for a BMP test assessment. They further suggested correcting the standard deviation of the samples considering the standard deviation of the inoculum. The experiment should be repeated if the coefficient of the variation of inoculum or the coefficient of a single sample is too high, and if the methane yield of the positive control is too high or too low.

In 2016, a revised version of the guideline VDI 4630 (VDI, 2016) was published with a stricter standard for batch tests. The standardization of BMP tests is very challenging and the main reason is that a BMP test is a biological test and the microbial growth and performance cannot be entirely defined.

2.4 Prediction of biogas and methane yields

Several authors have published studies on the mathematical modeling of anaerobic degradation and the prediction of biogas and methane yields. Two main groups of models exist within the literature: a) white-box approach, which is extremely demanding for modeling due to the high variability of microbes, feedstock specific microbial communities involved, and the different optimal growth conditions of the microbes (Mulka et al., 2016); b) black-box approach, which is a simple approach and allows the simulation of the biological process with less effort, however systematic effects (e.g., microclimate of local cultivation conditions, new cultivars, etc.) cannot be predicted by static models.

In this study only the black-box approach has been investigated as the objective was to create a global model for energy crops, which can be utilized by both the scientific community and practitioners.

The black-box approach can also be divided into two groups: stoichiometric and empirical models. A stoichiometric model refers to the chemical formula of the organic substrate, whereby the maximum methane (CH_4) and carbon dioxide (CO_2) production can be calculated based on an oxidation-reduction reaction involving water (Buswell and Mueller, 1952). These models do not consider any degree of biological degradation and produce similar results for various energy crops, since biomass always contains about 45 % of carbon. In order to determine the differences among the energy crops and to show reasonable results, advanced stoichiometric models were developed including chemical composition of the energy crops (Baserga, 1998). The chemical composition of the sample can be determined based on the laboratory analysis for animal feed (fodder analysis), using wet chemical analytical methods (classical laboratory analysis) or near-infrared spectroscopy (NIRS). Consequently, the share of the three main organic fractions carbohydrates, protein and fats can be defined. The percentage of each organic fraction can be multiplied with the reference value (stoichiometric biogas yield), respectively, to calculate the biogas yield of the sample. However, for energy crops this approach is very imprecise, since the difference between slowly digestible fractions such as cellulose and readily digestible fractions such as starch is not considered.

Empirical models were developed based on the chemical analysis of the samples and the experimental values of the batch tests for the biogas yield determination. Although, many studies have been performed, it still remains unclear what effect the different concentrations of the chemical compounds could have on the biogas yield. These models consider the difference between the slowly and readily digestible fractions, albeit insufficiently.

Table 1 shows a list of mathematical equations developed for the estimation of biogas yield. Buswell and Mueller (1952) developed a model that describes the stoichiometric reduction of the organic substrate. Boyle (1976) extended this model to include the chemical elements nitrogen (N) and sulphur (S), which are converted to ammonia (NH_3) and hydrogen sulphide (H_2S), respectively. The models of Buswell and Mueller (1952) and Boyle (1976) are based solely on the chemical formula of the organic substrate, and the biological degradation is not addressed.

Table 1: Selected models for biogas and methane yield prediction.

Author	Model
Buswell und Mueller (1952)	$C_aH_bO_c + \left(a - \frac{b}{4} - \frac{c}{2}\right) H_2O \rightarrow \left(\frac{a}{2} + \frac{b}{8} - \frac{c}{4}\right) CH_4 + \left(\frac{a}{2} - \frac{b}{8} + \frac{c}{4}\right) CO_2$
Boyle (1976)	$C_aH_bO_cN_dS_e + \left(a - \frac{b}{4} - \frac{c}{2} + \frac{3d}{4} + \frac{e}{2}\right) H_2O \rightarrow \left(\frac{a}{2} + \frac{b}{8} - \frac{c}{4} - \frac{3d}{8} - \frac{e}{4}\right) CH_4 + \left(\frac{a}{2} - \frac{b}{8} + \frac{c}{4} + \frac{3d}{8} + \frac{e}{4}\right) CO_2 + dNH_3 + eH_2S$
Baserga (1998)	Y_B = 790 (XF + NfE) + 700 XP + 1250 XL
Keymer and Schilcher (1999)	Y_B = 790 (XF VQ_{XF} + NfE VQ_{NfE}) + 700 XP VQ_{XP} + 1250 XL VQ_{XL}
Kaiser (2007)	Y_B = 307 XP + 781 XL + 627 OR + 938 HC + 691 CL − 358 ADL Y_M = 147 XP + 560 XL + 289 OR + 459 HC + 382 CL − 271 ADL
Weißbach (2008)	Y_B = 800 FOM Y_M = 420 FOM
Triolo et al. (2011)	Y_M = 447 – 7 CL – 277 ADL
Rath et al. (2013)	Y_B = 64.83 – 2678 ADL + 4299 HC + 12857XL – 3657 RS
Thomsen et al. (2014)	Y_M = 347 (CL + HC + R) – 438 ADL

The model of Baserga (1998) refers to the three main organic components: carbohydrates (crude fiber (XF) and nitrogen free extract (NfE)), crude protein (XP) and crude lipids (XL). The share of the sample's chemical composition can be defined by fodder analysis and the ultimate biogas yield of each chemical compound is calculated with the stoichiometric model (Table 2). However, this model does not consider the lignocellulosic matrix and 100 % degradability is always assumed. Due to this fact, the

biogas yield prediction is similar for most energy crops species and no difference can been observed for the cultivars of a species.

Table 2: Biogas yield and methane content of the three main organic components (Baserga, 1998).

Organic component	Biogas yield [L/kg$_{VS}$]	CH_4 content [%]
Carbohydrates	790	50
Crude lipids	1250	68
Crude protein	700	71

Keymer and Schilcher (1999) modified the model of Baserga (1998) by including the quotient of digestibility (VQ) from DLG (1997), which were determined empirically by biological methods (in vitro or in vivo). The model is based on the assumption that the microbiological degradation in the rumen is similar to degradation in the biogas plant. Although the biological process of the two systems (rumen and biogas plant) is similar, the final products (VFAs vs. biogas) of the systems and physical properties (e.g., retention time, particle size etc.) are different. This may cause the biogas yield prediction of energy crops to be underestimated.

Multiple linear regression models were also used for BMP prediction. Kaiser (2007) developed a model based on regression analysis among the share of the six components (crude protein, crude lipid, organic rest, hemicellulose, cellulose, and lignin) and the experimental values (BMP tests) of biogas yield from energy crops. Because the substrate on the models of Kaiser (2007) is divided into several components, a higher accuracy is expected. However, for external datasets the estimation error of biogas yield prediction is high, since the inter-correlation of the regressors is not considered and causes a high prediction error.

Weißbach (2008) developed a model based on the determination of the content of fermentable organic matter (FOM). FOM as a parameter is described by a regression function, with each energy crop needing a separate function. For these models, only the laboratory analyses of crude ash (XA) and XF are necessary. The models are based on the assumption that only the carbohydrates are the main source for biogas production, and the lipid and protein contents have only little influence on the biogas yield (Weißbach, 2009a).

Triolo et al. (2011) published regression models to assess biodegradability of energy crops and manure. In both cases lignin (ADL) and cellulose (CL) content were suitable

variables for BMP prediction. On the one hand, lignin is indigestible under anaerobic conditions and therefore is expected to have a low biogas yield. On the other hand, cellulose is well degradable under anaerobic conditions and often is used as a reference substrate in laboratory-scale experiments. Triolo et al. (2011) found that both variables (ADL and CL) contributed negatively to BMP prediction when both are used at the same time as regressors. This result of CL is contrary to the literature, since CL was expected to be positively correlated with biogas yield. However, this can be explained by the fact that all chemical compounds are inter-correlated and by the fact that the regressors also describe co-effects.

Rath et al. (2013) developed a maize-specific model in order to identify the differentiation among maize cultivars. The model shows that lignin and sugars negatively influence the biogas yield (Y_B), whereas crude lipids and hemicellulose have a positive influence. However the model is very specific and cannot be used for other energy crops.

Thomsen et al. (2014) developed a global model for energy crops and proved to be in agreement with Triolo et al. (2011) by showing that lignin is the most suitable variable for a mono-causal regression. There are two versions of this model based on the biomass analysis method. In this study the model based on fodder analysis was used. The authors used their own data, but also values from the literature for a total of 64 samples in order to develop a global model for BMP prediction. The model was presented as $Y_M = 347\ (CL + HC + R) - 438\ ADL$, but since residuals (R) were defined as $R = 1 - CL - HC - ADL$, the regressors of the model can be limited only to ADL, $Y_M = 347 - 785\ ADL$.

3 Problem statement and objectives

Biomass can be degraded by a biochemical process in the absence of oxygen (anaerobic digestion) and the products of this process are digestate and biogas (mainly CH_4 and CO_2). The biogas and biomethane potential (BMP) are important parameters for assessing any biodegradable material in order to define its suitability as feedstock in a biogas plant. However, the experimental determination of the BMP is a time-consuming and costly process. The fodder analysis is a standardized method to assess the nutritive value of feedstock; these values can be taken from pre-existing databases or the samples can be analyzed with low cost and readily available test kits. Therefore, a BMP prediction based on the chemical composition of the feedstock is desirable.

In previous studies, the lignin content was strongly negatively correlated to the biogas and methane yields (Thomsen et al., 2014; Triolo et al., 2011). Simple mono-causal regression models were developed with lignin being the only regressor for biogas and methane yield prediction. Nevertheless, the accuracy in the mono-causal models is low and multiple regression is required (Gunaseelan, 2009; Xu et al., 2014). At this point, it should be noted that the regressors' selection is limited due to inter-correlation among the chemical compounds of the feedstock. Various authors have developed models to predict the biogas and methane yield from energy crops based on fodder analysis. However, the effect of the chemical compounds on the biogas production cannot be clearly determined. Among the studies, contradictory results were published considering the effect of each chemical compound on the methane yield. Moreover, the high number and the selection of the regressors led to poor performance of the models with an external dataset.

The partly contradictory results published so far and the lack of studies which are aimed to address the diversity of classical agricultural plants motivated this study. The objectives of this study were: Firstly, to analyze the correlations and interactions of the chemical compounds with the biogas and methane yields and to develop a model to estimate the biogas and methane yields across different plant species; secondly, to develop a specific model that predicts the biogas potential of grassland samples focusing on the question whether a more precise prediction of the biogas yield relies on a strict selection of samples than presented so far; thirdly, to analyze the correlation between the hydrolysis rate constant and the chemical composition of the feedstock based on a high variety of plant species and cultivars and to develop and evaluate a regression model in order to predict hydrolysis rate constant; and fourthly, to assess and validate the developed prediction models and to compare them with previously published ones.

4 Research questions and hypotheses

The following research questions are introduced and addressed:

1. How are the biogas and methane yields affected by the chemical composition of energy crops?
2. Is the effect of a specific chemical compound on the biogas production independent of the plant species?
3. Can the respective effect be expressed as a continuous function?
4. Which chemical components cause a statistically significant change on the biogas production and which of them can be used simultaneously in a prediction model?
5. Can the biogas production rate in a BMP test be accurately predicted?

At this point it should be clarified that firstly, a statistically significant effect of a chemical compound on biogas production is not necessarily connected with its properties, because the relationship between the compound and biogas yield is described only empirically and not by the biochemical degradation process steps. For instance, the structure of a protein and the various protein groups cannot be considered since all protein groups are recorded as crude protein based on the nitrogen content. Secondly, the parameters of the fodder analysis are analytical parameters and do not necessarily describe the biological and chemical structure of the whole plant (Jung, 1997). Thirdly, one of the assumptions when performing multiple linear regression analysis is that the regressors are not inter-correlated. The parameters of the fodder analysis are expressed in percentage of total solids, and the share of one compound will automatically change if the content of another changes. Hence, co-effects between the regressors cannot be avoided.

Based on the research questions and with regard to the analytical *status quo*, the following hypotheses were proposed and tested:

a.) The variation in biogas yield is a function of the plant's chemical composition (**Paper I**).
b.) The biogas potential can be mathematically described by the plant's chemical components (**Paper I**).
c.) For high estimation accuracy of biogas yield, plant groups or even plant species, specific models are needed (**Paper II**).
d.) More than two regressors are needed for high accuracy and carbohydrates (fiber and not) alone cannot predict accurately biogas yield (**Paper II**).
e.) The hydrolysis rate constant can be predicted based on the plant's chemical compounds (**Paper III**).
f.) The determination of the hydrolysis rate constant affects its correlation with the chemical compounds (**Paper III**).

g.) The regression model can precisely predict each individual sample of an independent dataset (**Paper IV**).
h.) Linear regression models are suitable for feedstock assessment (**Paper IV**).

To test these hypotheses, field experiments were carried out in order to collect different plant species and cultivars of energy crops under well-defined growing conditions. Samples were analyzed according to the Weender and van Soest methods. The BMP tests were run with a high temporal resolution in order to identify statistical correlations between the chemical composition of the samples and their biogas potential.

5 Materials and methods

5.1 Test substrates

The various species and cultivars of energy crops used in this experiment were grown in field experiments under defined conditions and were harvested depending on their development according to the Biologische Bundesanstalt, Bundessortenamt und Chemische Industrie (code of the phenological growth stages of plants) (BBCH)-scale (Hess et al., 1997; Meier, 2018). The BBCH scale is a system for coding phenologically similar growth stages of plant species. It is a decimal system, with 10 principal growth stages and up to 10 secondary ones. The principal and the secondary growth stages are descripted with a two-digit code.

The field experiments were performed from 2012 to 2016. After harvest, the samples were immediately dried in a convection oven at 40 °C to stabilize the actual plant status. The samples were ground to 10 mm with a cutting mill (RETSCH, model SM 200, Haan, Germany) in order to achieve a homogeneous sample. After that, they were stored at room temperature. Fodder analysis and a BMP test were performed for all samples. The values (biogas yield and chemical composition) of each sample are given in Appendix (Paper I, II, III and IV). The selected samples can be divided into four groups:

- Various energy crops at individually defined optimal harvest times, which are then suitable to be used as feedstock in a biogas plant (harvested 2012; 41 samples).
- Selected species of grassland systematically harvested at different developmental stages within certain re-growth periods and in two different years (harvested 2013 and 2014; 116 samples).
- Selected maize cultivars systematically harvested at specific developmental stages to produce maize stover, which is suitable to be used as feedstock in a biogas plant (harvested 2014; 64 samples).
- Selected perennial energy crops systematically harvested at different developmental stages (harvested 2015; 91 samples).

5.1.1 Various energy crops

The first set of samples was selected to determine the global influence of the chemical composition of various plant species on the biogas yield, 41 samples of 11 different plant species were investigated (Table 3). The samples were selected according to the following criteria: a) the samples should be suitable feedstock for a biogas plant in Bavaria; b) the samples should be harvested at a BBCH-code, when a reasonable

amount of biomass can be selected, and c) the biogas and methane yields of the samples should follow a normal distribution.

Table 3: Number of samples within the 11 investigated plant groups of the first set of samples.

Samples	**n**
Cup plant	2
Barley	2
Grassland	9
Millet	4
Potato	2
Clover	3
Maize	4
Rye	4
Sunflower	2
Triticale	8
Sugar beet	1

5.1.2 Grassland biomass

The second set of samples consisted of selected grassland species. The species selection and the field experiments were performed by the Institute for Crop Science and Plant Breeding (IPZ) at the Bavarian State Research Center for Agriculture (LfL) in Freising, Germany. The most important criterion for the selection of plant species and harvest dates was their use of grassland as feedstock for an agricultural biogas plant.

Two field experiments were conducted and six selected grassland plant species were investigated (Table 4). Four grass species (*Lolium perenne*, *Dactylis glomerata*, *Poa pratensis*, *Festuca pratensis*) and two legume species (*Trifolium pratense*, *Trifolium repens*) were grown in field plots (10 m^2). In addition, four cultivars of *Lolium perenne* (Table 4) were tested. Each species of plant was grown on a single field plot.

Table 4: Investigated grassland plant species of the second set of samples.

Species	**Cultivar**	**n**
1. Ray grass (*Lolium perenne)*	1a. Arvicola	12
	1b. Respect	14
	1c. Sponsor	10
	1d. Sirius	11
2. Orchard grass (*Dactylis glomerata*)	2. Husar	14
3. Common meadow-grass (*Poa pratensis)*	3. Lato	12
4. Meadow fescue (*Festuca pratensis)*	4. Preval	13
5. Red clover (*Trifolium pratense)*	5. Titus	15
6. White clover (*Trifolium repens)*	6. Lirepa	15

The selected grass and clover species were grown under defined conditions at the test site in Pulling, Germany during 2013 and 2014. To obtain information about the change of the chemical composition of the plants during their development, each plant underwent several harvests. The goal was to create a dataset with a wide range of chemical compositions, which would also be suitable for statistical analysis.

The collection of the samples took place during the first, second, and third growth cycles with advancing harvest dates in each growth cycle. At defined phenological development stages based on the BBCH scale, five harvest dates were scheduled during the first growth cycle and three harvest dates were scheduled during each the second and third growth cycles. However, due to the weather conditions the plants were not harvested according to the schedule (Table 5, Table 6).

Table 5: Harvest dates of the investigated grassland species in 2013. Numbers in the parentheses denote the development stage according to the BBCH-Code.

Species	Cultivar	1. Growth cycle					2. Growth cycle			3. Growth cycle		
		1. HD	2. HD	3. HD	4. HD	5. HD	1. HD	2. HD	3. HD	1. HD	2. HD	3. HD
Lolium perenne	Arvicola	02. May (41)	06. May (49)	08. May (52)	14. May (55)	06. Jun. (65)	16. Jul. (30)	-	-	14. Sep. (29)	-	-
	Respect	14. May (43)	16. May (49)	17. May (51)	21. May (55)	06. Jun. (60)	18. Jul. (30)	-	-	06. Sep. (29)	-	-
	Sponsor	19. May (39)	-	12. Jun. (53)	15. Jun. (56)	18. Jun. (60)	-	-	-	-	-	-
	Sirius	19. May (39)	-	12. Jun. (53)	15. Jun. (56)	18. Jun. (60)	-	-	-	-	-	-
Dactylis glomerata	Husar	14. May (41)	16. May (45)	19. May (51)	21. May (56)	13. Jun. (60)	16. Jul. (30)	-	-	14. Sep. (29)	-	-
Poa pratensis	Lato	06. May (41)	08. May (47)	14. May (52)	17. May (56)	06. Jun. (61)	23. Jul. (30)	-	-	04. Sep. (30)	-	-
Festuca pratensis	Preval	14. May (41)	16. May (45)	19. May (51)	21. May (55)	13. Jun. (60)	11. Jul. (30)	-	-	14. Sep. (30)	-	-
Trifolium pratense	Titus	16. May (51)	17. Jun. (55)	20. Jun. (61)	21. Jun. (65)	-	26. Jul. (55)	02. Aug. (61)	-	29. Aug. (59)	05. Sep. (65)	14. Sep. (69)
Trifolium repens	Lirepa	16. May (51)	12. Jun. (55)	15. Jun. (59)	17. Jun. (65)	-	05. Jul. (61)	16. Jul. (69)	23. Jul. (69)	29. Aug. (61)	05. Sep. (69)	-

HD: Harvest date

Table 6: Harvest dates of the investigated grassland species in 2014. Numbers in the parentheses denote the development stage according to the BBCH-Code.

Species	Cultivar	1. Growth cycle					2. Growth cycle			3. Growth cycle		
		1. HD	2. HD	3. HD	4. HD	5. HD	1. HD	2. HD	3. HD	1. HD	2. HD	3. HD
Lolium perenne	Arvicola	-	-	06. May (51/53)	14. May (55/59)	22. May (65)	-	-	25. Jul. (43/65)	09. Oct. (32/33)	-	-
	Respect	08. May (39)	-	14. May (52)	-	02. Jun. (60)	06. Jun. (42)	17. Jun. (57/59)	24. Jun. (65)	09. Oct. (32/33)	-	-
	Sponsor	14. May (37)	02. Jun. (52)	-	06. Jun. (55)	18. Jun. (59/60)	-	-	17. Jul. (61/65)	09. Oct. (32/33)	-	-
	Sirius	14. May (37/38)	22. May (40)	02. Jun. (51)	06. Jun. (57)	18. Jun. (59/60)	-	-	17. Jul. (61/65)	09. Oct. (32/33)	-	-
Dactylis glomerata	Husar	06. May (39)	14. May (42)	-	22. May (53/55)	06. Jun. (59/60)	23. Jul. (43)	22. Aug. (47)	-	09. Oct. (32/33)	-	-
Poa pratensis	Lato	-	-	-	06. May (55)	26. May (61)	25. Jul. (39)	02. Sep. (47)	-	09. Oct. (32/33)	-	-
Festuca pratensis	Preval	08. May (38)	-	14. May (51)	22. May (57)	26. May (59/60)	23. Jul. (43)	-	-	09. Oct. (32/33)	-	-
Trifolium pratense	Titus	03. Jun. (51)	-	17. Jun. (57)	24. Jun. (65/69)	-	04. Jul. (55)	18. Jul. (60/61)	23. Jun. (63)	-	-	-
Trifolium repens	Lirepa	26. May (51)	02. Jun. (55)	06. Jun. (59)	17. Jun. (65/69)	-	-	-	24. Jun. (69)	-	-	17. Jul. (61/65)

HD: Harvest date

Figure 2 represents the average monthly precipitation across ten years (2003-2012), as well as the monthly precipitation in both 2013 and 2014 per month. The data was obtained from the weather station in Freising, Germany. The harvests took place in May/June, in July/August and in September/October, for the first, second and third growth cycle, respectively, as documented in Tables 5 and 6. Water availability is crucial for the development (quality) and growth (quantity) of grasslands. In May, September and October in both 2013 and 2014, the monthly values of precipitation were higher than the ten-year average. In June 2013 precipitation was also higher, but was extremely low in July 2013 and vice versa in 2014.

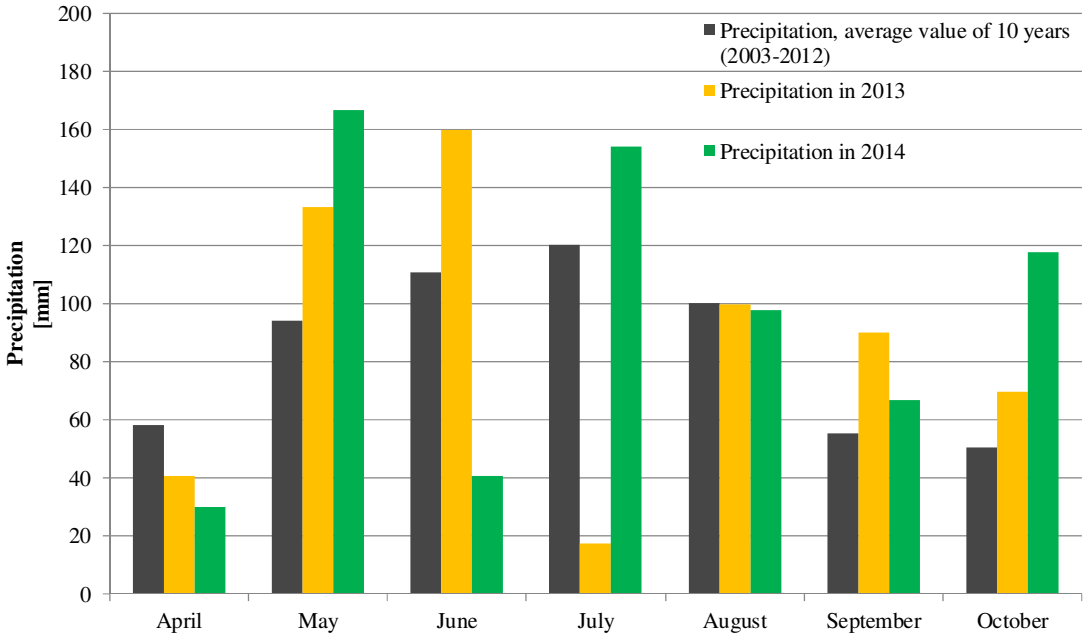

Figure 2: Precipitation observations at the weather station in Freising, Germany.

The first harvest of the first growth cycle is the most important since that is when the highest biomass amount can be collected. In contrast, the second growth cycle usually has a lower biomass production and its composition is more variable. The third growth cycle is characterized by very low biomass production, but with a rather repeatable plant composition. For this reason, various cuts (harvest dates) of each growth cycle were investigated.

5.1.3 Maize stover

The third set of samples consisted of four different maize cultivars. The maize was harvested at full maturity, i.e. BBCH ≥ 89, since the goal was to create a dataset containing agricultural residues, i.e. maize stover. Three different harvest dates were undertaken two in October and one in November. Furthermore, each fraction of the maize stover (leaves, stalks, cobs) was tested separately. The species selection and the field experiments were performed by the Institute for Crop Science and Plant Breeding (IPZ) of the Bavarian State Research Center for Agriculture (LfL) in Freising, Germany.

5.1.4 Perennial energy crops

The fourth set of samples consisted of eight perennial crops and two annual cultures (maize and rye) (Table 7). The samples were grown in various locations in the state of Bavaria, Germany and harvested at different harvest dates or cycles. The species selection and the field experiments were performed by the Technology and Support Centre (TFZ) in Straubing, Germany.

Table 7: Investigated plant species of the fourth set of samples.

Common name	**Species**	**n**
Maize	*Zea mays*	6
Rye	*Secale cereale*	5
Wild rye	*Secale multicaule*	14
Virginia Wildrye	*Elymus virginicus*	2
Sorghum	*Sorghum bicolor x Sorghum sudanense*	1
Tall wheatgrass	*Elymus elongatus*	30
Reed canary grass	*Phalaris arundinacea*	1
Virginia mallow	*Sida hermaphrodita*	11
Cup plant	*Silphium perfoliatum*	8
Switchgrass	*Panicum virgatum*	13

5.2 Experimental set-up

In total, 19 incubators with 228 digesters were available for the BMP tests (Fig. 3). The BMP tests were performed according to the guideline VDI 4630 (VDI, 2006) and the VDLUFA method book (VDLUFA, 2011). Each incubator contains 12 digesters; each digester is connected by tubes with milligascounters (Ritter Apparatebau GmbH, Bochum, Germany) for gas volume recording. A gas bag was attached to three digesters (three technical replicates) and a gas analysis was performed for every 1.5 L biogas produced. The batch digesters have a total volume of 2 L and a working volume of approximately 1.4 L. The experiments were run at mesophilic conditions, i.e. the temperature was set to 38 ±1 °C. Two control samples were included to verify the biological activity of the inoculum. Microcrystalline cellulose and a defined sample of dried whole plant maize served as positive controls. Additionally, the inoculum alone was tested to determine its gas potential.

Figure 3: Batch system of LfL for BMP tests.

Each sample was tested in triplicate (technical replicate) (Fig. 4). Microcrystalline cellulose was used as a reference sample, as well as for correcting the biogas yield for all experiments. In order to ensure statistical accuracy, six replicates (two analytical replicates) were used with cellulose. For dry samples, each digester was filled with 400 mL of distilled water, 1000 g fresh matter (FM) inoculums, and 20 g of FM sample. The ratio of volatile solids of the test substrate to volatile solids of the inoculum was

0.5 ± 0.1. The TS content in the digester was between 4 and 5 % of FM. If the daily biogas production was less than 0.5 % of the total volume of biogas produced, the measurement was terminated.

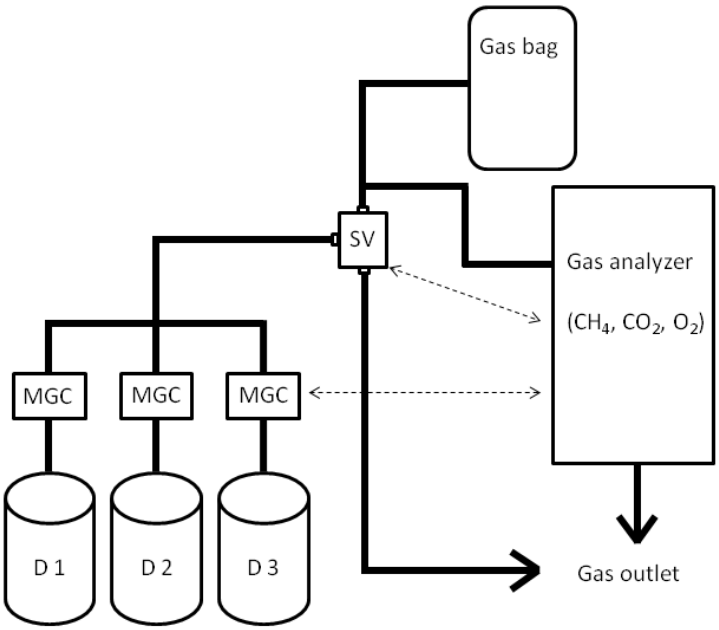

Figure 4: Schematic diagram of the BMP test system of LfL. D: digester, MGC: Milligascounter, SV: solenoid valve.

The volume of the produced biogas was measured with an accuracy of ± 3 %. The produced biogas flows through the gas inlet nozzle into a capillary of the counter tank, which is filled with silicon oil. The gas then moves to the two-chamber measurement cell. The measuring of the gas volume occurs in discrete steps by counting the tilts of the measurement cell with a resolution of approximately 1 mL. Each gas counter was regularly calibrated to define the exact volume for each tilt. The measurements were recorded online and the data was stored on an hourly basis.

During the experiment, around 33 gas analyses took place per sample for the biogas produced from the three digesters. The gas analysis was carried out using infrared sensors for the content of methane and carbon dioxide, and an electrochemical sensor for the content of oxygen (Awite Bioenergie GmbH, Langenbach, Germany). Table 8 shows the range and the accuracy for each individual analysis unit. Air pressure and temperature were recorded each hour. The saturation vapor pressure in the biogas was calculated by the Magnus formula and the biogas volume was normalized at 273.15 K and 1013.25 hPa. The biogas yields are reported as per liter of dry gas at standard temperature and pressure per kilogram volatile solids added (L/kg_{VS}). The value of the

biogas yield was the average value of the three replicates with a coefficient of variation (CV) less than 10 % (Heuwinkel et al., 2009).

Table 8: Overview of the measuring principle, range, repeatability and accuracy of the sensors of the gas analyzer.

Sensor	Principle	Range	Repeatability	Accuracy
CH_4	Infrared	0 – 100 Vol.- %	± 0.2 %	± 2 %
CO_2	Infrared	0 – 100 Vol.- %	± 0.2 %	± 2 %
O_2	Electrochemical	0 – 25 Vol.- %	± 0.1 %	± 1 %

5.3 Inoculum

The origin of the inoculum and its properties is very important for the BMP experiments, since the microbial community can influence the biological degradability of the substrate (Gu et al., 2014; Lopes et al., 2004; Xu et al., 2014). Therefore, inocula with similar chemical and physical properties were used for all BMP experiments performed in this study.

To obtain a defined biocoenosis, a pilot digester (continuously stirred-tank reactor) with a working volume of 2.5 m^3 was operated under steady-state conditions (the coefficient of variation of 5 days of methane productivity was continuously less than 10 %). The digester was run at an organic loading rate of 3.0 $kg_{VS}/(m^3{*}d)$ with an 80 % cattle manure and 20 % dairy cattle feed mixture (total mixed ration (TMR)). TMR was composed of 44.4 % maize silage, 39.5 % grass silage, 4.9 % supplementary feed (23.4 % shredded barley, 23.4 % shredded corn, 46.9 % grain maize, 1.1 % calcium carbonate, 0.3 % cattle salt and 4.8 % VitalMiral Hofmix (RKWSüd)), 3.7 % hay, 3.7 % molasses, 2.5 % Bovigold ® SojaPlus (BayWa) and 1.2 % straw. The digester was operated at 38 ± 1 °C and the hydraulic retention time (HRT) was 19 days. The digester was located in Freising, Germany.

To prove the current fitness of the inoculum, chemical analyses were carried out regularly. One week prior to BMP testing, the effluent of the digester (defined biocoenosis) was sieved through a 10 mm sieve and was stored at the test temperature of 38 ± 1 °C without feeding to reduce endogenous biogas potential. The degassed material was used as an inoculum for the BMP experiments. Table 9 demonstrates the average values of the chemical parameters of the inoculum.

Table 9: Chemical characterization of the inoculum just prior to the experiments (n = 17).

Parameter	Unit	Average value
TS	[% FM]	4.37 ±0.25
VS	[% FM]	3.12 ±0.24
pH	[-]	7.7 ±0.1
TIC	[mg/kg$_{FM}$]	8,851 ±625
VFA	[mg/kg$_{FM}$]	256 ±92
VOA/TIC	[-]	0.18 ±0.05
NH_4^+-N	[mg/kg$_{FM}$]	1,637 ±144

5.4 Chemical analysis

All samples were analyzed for their chemical composition. In addition, the inoculum of the BMP test was analyzed before and after each experiment. Table 10 presents the methods that were used to determine the laboratory parameters. Fodder analyses were performed according to the methods of the Association of German Agricultural Analytic and Research Institutes (VDLUFA, 1976). All chemical analyses were carried out in duplicate by the Central Department for Quality Assurance and Analytics (AQU) at LfL in Freising, Germany.

According to the fodder analysis methods, some of the parameters were analyzed and some were calculated, but all were consistently expressed as percentages. Total solids (TS), crude ash (XA), crude lipid (XL), crude fiber (XF), starch (ST), reducing sugar (RS), neutral detergent fiber (NDF), acid detergent fiber (ADF) and acid detergent lignin (ADL) were measured. Volatile solids (VS), crude protein (XP), nitrogen free extract (NfE), hemicellulose (HC), cellulose (CL), non-fiber carbohydrate (NFC), and organic residue (OR) were calculated as follows:

$$VS = TS - (XA \times TS / 100),$$

$$XP = 6.25 \times N,$$

$$NfE = 100 - XA - XP - XL - XF,$$

$$HC = NDF - ADF,$$

CL = ADF – ADL,

NFC = 100 – XA – XP – XL – NDF,

OR = 100 – XA – XP – XL – ST – RS – NDF.

TS and VS content are expressed as % of FM, all others parameters are expressed as % of TS. However, due to the fact that only the organic complex can be used during the anaerobic digestion, the contents of all chemical compounds were expressed as grams per kilogram of volatile solids added (g/kg_{VS}).

Table 10: Parameters and methods of chemical analyses.

Parameter	**Guideline**	**Remarks**
Total solids (TS)	DIN 12880 DIN 38414-S2	Dry the sample at 105 °C
Crude ash (XA)	DIN 12879 DIN 38414-S3	Burn the sample at 550 °C
pH	DIN 38404-C5 DIN 12176	Glass-Electrode
Total inorganic carbon (TIC)	DIN 38409 H7	Titration
Volatile organic acids to total inorganic carbon (VOA/TIC)	Titrator, TitroLine® alpha plus (SCHOTT AG, Germany)	Titration pH 5.0 and 4.4
Total volatile fatty acids (VFA_{total})	DIN 38414 S.19	Distillation and titration
C2-C8 volatile fatty acids (VFA_{GC})	Agilent 6890N Gas chromatograph (Agilent Technologies, USA)	Gas chromatography
Ammonium nitrogen (NH_4^+-N)	DIN 38406 E5-2	Distillation and titration
Crude protein (XP)	VDLUFA MB III 4.1.2	Dumas-Method
Crude lipid (XL)	VDLUFA MB III 5.1.1/2	

Crude fiber (XF)	VDLUFA MB III 6.1.1/2	Weender - Method
Reducing sugar (RS)	VDLUFA MB III 7.1.1	Volumetrically
Starch (ST)	VDLUFA MB III 7.2.1	Polarimetric Method
Neutral detergent fiber (NDF)	VDLUFA MB III 6.5.1	Van Soest - Method
Acid detergent fiber (ADF)	VDLUFA MB III 6.5.2	Van Soest - Method
Acid detergent lignin (ADL)	VDLUFA MB III 6.5.3	Van Soest - Method

5.5 Statistical analysis

The correlation among the parameters was tested by the Pearson correlation coefficient (r) and a cross correlation matrix was developed to show inter correlations among the parameters. The normal distribution and the residues of the regression were analyzed using the Kolmogorov-Smirnov test. Furthermore, descriptive statistics such as standard deviation, coefficient of variation, standard error were conducted in order to compare and evaluate the datasets.

In order to reduce the interrelated effect of the parameters, while keeping the variation as high as possible, principal components analysis (PCA) was used. PCA can expose and visualize correlations within the dataset. Similarities and differences among the samples can be revealed by plotting the measured parameters in a plot of the mathematically defined variables called principal components (PC). Parameters that are correlated with each other but are largely independent from other subsets of parameters are combined into principal components. The PCs are the linear combination of observed parameters and are not correlated (orthogonal) with each other. The first PC extracts the maximum variability of the observed parameters. Sequential PCs are formed from the correlation of the residuals and are orthogonal to all other PCs (Jolliffe, 2002; Tabachnick and Fidell, 2012).

In addition, regression analysis was performed to develop models for biogas yield prediction. Moreover, the Mallows' Cp model selection method has been applied in order to find out the best fit model. A low Cp value indicates a good fit and robustness of the model. The parameters, coefficients of determination (R^2), the root mean square

error (RMSE), the coefficient of variation of the RMSE (CVRMSE) and the systematic error (bias) were used for evaluation and comparison of the models.

The software 9.3 SAS (SAS Institute, USA) and Unscrambler 10.3 (CAMO Software, Norway) were used for statistical analysis.

6 Research accomplishments

6.1 Paper I

Correlation between biogas yield and chemical composition of energy crops. ***Bioresource Technology 174, 2014, 316–320.***
https://doi.org/10.1016/j.biortech.2014.10.019

The objectives of the first paper were firstly to investigate the influence of the chemically defined compounds of various energy crops on biogas and methane yields, and secondly to develop a global model for BMP prediction.

In total, 41 different plants were analyzed in batch and their composition was chemically determined. A wide range in Y_B and Y_M was recorded; the Y_B ranged from 339 to 799 L/kg_{VS} and the Y_M from 177 to 401 L/$kg_{VS.}$ For ADL content below 10 % of TS, a strong negative correlation between ADL and BMP was recorded. Based on a simple regression analysis, more than 80 % of the variation of the samples can be explained through ADL. According to the PCA, it is concluded that only carbohydrates (fiber and non-fiber) are mathematically suitable variables for BMP prediction, even though XP and XL are characterized by a higher BMP than carbohydrates. This finding is in line with the observation of Weißbach (2008) and can be explained by the fact that no protein or fat rich samples were dominate in this dataset.

Statistically, it was shown that HC and ADL were suitable regressors for biogas yield prediction across plant species. This finding is in line with Triolo et al. (2011) and Thomsen et al. (2014), who also showed that ADL has a negative effect on BMP and the main influence on BMP prediction. Furthermore, Rath et al. (2013) also reported a strong positive effect of HC on BMP.

Based on the results of this paper both hypotheses (see Chapter 4) have been tested and it can be confirmed that a) the variation in biogas yield is a function of the plant's chemical composition and b) the biogas potential can be mathematically described by the feedstock's chemical compounds.

6.2 Paper II

Correlation between biogas yield and chemical composition of grassland plant species.
Energy Fuels 29, 2015, 7221–7229.
https://doi.org/10.1021/acs.energyfuels.5b01257

The objectives of the second paper were firstly to investigate the accuracy of Y_B prediction of energy crops using a global model, and secondly, to develop a specific model that predicts the biogas potential of grassland samples. Moreover, this paper focused on whether a precise prediction of biogas yield relies on a more strict selection of samples.

This study investigated and analyzed the composition and the Y_B of different grassland plant species at various harvest dates during the first three growth cycles to investigate the accuracy of Y_B prediction.

According to the batch trials, the Y_B ranged from 500 to 768 L/kg_{VS} and the Y_M from 263 to 425 L/kg_{VS} with a coefficient of variation of approximately 10 %. Although all tested samples are grassland species, they can be classified into two groups based on the plant family: The first group is the grass species and the second group is the legume species. A t-test revealed that the average XP and ADL contents of grass species were significantly lower than those of the legumes species. Moreover, the average RS, HC and CL contents of grass species were significantly higher than those of the legumes species. However, no significant difference was found in the average XL content between the two plant groups.

While the results proved that a global model is not sufficient to predict accurately the Y_B of typical grassland species, the grassland specific model increased the accuracy of BMP prediction. In particular, the developed grassland model reduced the estimation error to 5 %. The present model predicts the biogas yield of grassland plant species with an accuracy of 31 L/kg_{VS} using three regressors, namely ADL, HC and XP content. According to the model, ADL reduces biogas yield, while both HC and XP increase it. The regressors of the grassland model reflect the necessity to describe the actual physiological status of the plant.

In conclusion, it was shown that plant-group species specific models can predict the BMP with high estimation accuracy. However, more than two regressors are needed and crude protein was needed in addition to carbohydrates for accurate BMP prediction. Hence, both hypotheses (see Chapter 4) of this paper have been accepted.

6.3 Paper III

Correlation between hydrolysis rate constant and chemical composition of energy crops.
Renewable Energy 118, 2018, 34–42.
https://doi.org/10.1016/j.renene.2017.10.100

The objectives of the third paper were firstly to analyze the correlation between the hydrolysis rate constant (k_h) and the chemical composition of the feedstock based on a high variety of plant species and cultivars, and secondly, to develop and evaluate a regression model in order to predict k_h based on the feedstock chemical composition.

Fodder analysis and BMP tests with high temporal resolution were performed in order to identify statistical correlations between the hydrolysis rate constant (k_h) and the chemical composition of various energy crops. Different species and cultivars of energy crops were analyzed in order to develop a broadly applicable regression model for the prediction of k_h.

The hydrolysis rate constant was defined at 50 % of Y_B produced as follows:

$$k_{h\_0.5} = \frac{\ln(2)}{t_{0.5}}$$

The results indicated that the analytical parameters NFC and XP were statistically suitable for a multiple linear regression model for the prediction of k_h. In addition, the k_h prediction model was combined with the biogas yield prediction model presented in **Paper I** in order to predict both the biogas yield and biogas production rate based on a first-order kinetic model. Finally, the modeling approach was validated by an independent dataset. The results indicated that a first-order model can reproduce the biogas production rate during a BMP test and linear regression models can precisely predict the differentiation of the biogas production of various energy crops. The proposed approach offers a fast and reliable prediction of the biogas production rate and allows a feedstock assessment according to their biogas potential.

The hypotheses of this paper (see Chapter 4) have been accepted and it was confirmed that a) the hydrolysis rate constant can be predicted based on the plant's chemical compounds, namely non-fiber carbohydrates and crude protein, and b) the hydrolysis rate constant, determined at the time period of the half-maximum biogas production ($k_{h\_0.5}$), describes the kinetics of the biogas production in a BMP test well.

6.4 Paper IV

Predicting methane yield by linear regression models: A validation study for grassland biomass.
Bioresource Technology 265, 2018, 372–379.
https://doi.org/10.1016/j.biortech.2018.06.030

The objectives of this study were firstly, to assess and validate previously published prediction models with an independent dataset and secondly, to expose the power and limitation of linear regression models for biomethane potential prediction. Three global models for the methane yield prediction of energy crops and one grassland species specific model were assessed and validated. Two datasets were used for the validation, one with 55 individual samples of grassland species and one with average values from 9 cultivars.

The results revealed a similar performance of all four models for the individual samples. The correlation between the measured and predicted values of BMP was moderate and the prediction error 11 %. The models could not explain more than 27 % of the data variation of the individual samples. For the methane yield prediction of the average values, all four models performed well. Moreover, the grassland specific model represented the variation of the dataset with a correlation coefficient of 0.92 and achieved a prediction error of only 2 %. Hence, linear regression models are suitable in order to depict the variation of the BMP and to define a ranking of substrates. In order to minimize the prediction error and improve the estimation of the differences among the samples, average values of the same cultivar should be used. However, the prediction error for the absolute values may be high since systematic external effects cannot be identified from the regression models; therefore, the calibration dataset should always be updated.

The first hypothesis of this paper (see Chapter 4) cannot be entirely accepted. On the one hand, the regression model can precisely predict an independent dataset if similar information is included in the calibration dataset. On the other hand, an independent model validation exposes the insufficiency of a static model to explain external effects, such as new cultivars. The second hypothesis has been accepted and it was confirmed that linear regression models are suitable for BMP test prediction within the calibration range and this information can be used for feedstock assessment.

6.5 Summary of outcomes

In Chapter 4 the research questions and the hypotheses of this dissertation were stated. According to the outcomes of the four peer-reviewed published papers, the questions can be answered as follows:

1. How are the biogas and methane yields affected by the chemical composition of energy crops?

ADL has a clear negative impact on biogas yield with a correlation coefficient up to -0.9. Therefore, ADL seems to be the only regressor candidate for a mono-causal regression model of BMP prediction. The ADL content of energy crops is usually below 6 %, but the effect on BMP prediction is the highest. Lignin alone is likely not the sole contributor to a decrease in the BMP, as the complex lignocellulosic structure is proven to as well. Thus, the regressor ADL in the models reflects the negative influence of the whole fiber matrix. The HC was used in MLR to explain the positive effect of the fiber matrix on biogas production. For grassland species XP content significantly affected the biogas production. Although XP is occasionally negatively correlated with Y_B, it has a positive impact on Y_B, according to the MLR. This negative mono-correlation between XP and Y_B was observed probably because the protein content decreases during the plant senescence, while the fiber and ADL content increases, attributing to a lower digestibility. Hence, XP depicts plant maturity and it has a positive effect on Y_B. Statistically, for XL has not found any significant effect on biogas production; however it can be used for a species specific model (Rath et al., 2015). NFC and XP have been identified to be important variables for the prediction of $k_{h\_0.5}$. Probably because the variable NFC describes the share of all carbohydrates in the crop and carbohydrates (fiber and non-fiber) typically occupy more than 80 % of the total VS of energy crops and hence, are the main source for biogas production. XP was also statistically significant for the prediction model, since it characterizes the ageing of the crop, whereas young plants are characterized by high share of XP and good anaerobic digestibility and vice versa.

2. Is the effect of a specific chemical compound on the biogas production independent of the plant species?

Based on the regression coefficient of a specific chemical compound (e.g. ADL) on different regression models (local or global), it has been revealed that the impact of a specific chemical compound is dependent on the plant species, since different values of the regression coefficient are used. The matrix and the chemical structure of the compounds are plant species specific.

3. Can the respective effect be expressed as a continuous function?

It has been shown that the effect of ADL on BMP is a continuous function for ADL values up to 10 % of TS. Probably for values above 10 % of TS, the effect can no longer be descripted by a linear model. Triolo et al. (2012) characterized the ADL content of 10 % as a critical limit for biodegradability as well.

4. Which chemical components cause a statistically significant change on the biogas production and which of them can be used simultaneously in a prediction model?

The variables for a MLR should be as independent as possible, which means variables that summarize similar information of the dataset variation should not be used simultaneously. This study showed that no more than two variables of carbohydrates (fiber and non-fiber) can be used simultaneously. For a global BMP prediction model, it was shown that ADL and HC were suitable regressors. The prediction accuracy was improved for a grassland species model with the introduction of one more regressor, namely XP.

5. Can the biogas production rate in a BMP test be accurately predicted?

This study showed that a first-order model can simulate the biogas production rate in a BMP test, and linear regression models can precisely predict the variation on the biogas production of different energy crops. This approach can be used to assess the suitability of different feedstocks for biogas production.

7 Discussion and conclusions

7.1 Accuracy of the prediction

7.1.1 Across plant species model

In Chapter 6.1, a global biogas yield prediction model for energy crop species was introduced. The model was developed based on a dataset of 31 samples. Although the dataset was small, it can be considered representative. The samples were suitable as feedstock of a biogas plant; they were harvested at developmental stages, at which a reasonable biomass amount could be recovered. Moreover, the biogas yields of this dataset were normally distributed, which is an important requirement for further statistical analyses. However, the accuracy and the suitability of a regression model can be validated only with external datasets. The across plant species model was tested and validated with four independent external datasets: firstly, with a small dataset of 10 various energy crop samples similar to the dataset of the calibration (Chapter 6.1), secondly, with a medium dataset of 61 grassland species samples (Chapter 6.2), thirdly, with a large dataset of 131 various energy crops samples (Chapter 6.3), and fourthly, with a large dataset of 91 perennial energy crops samples (Chapter 6.3). These four validation approaches of the across plant species model (global model) can define the potential application areas of the model and the boundaries of its application range. In Table 11, the results of these three validations are summarized.

Table 11: Performance of the across plant species model tested with different independent datasets.

Parameter	**1. Dataset**	**2. Dataset**	**3. Dataset**	**4. Dataset**
Number of samples	10	61	131	91
Slope	1.00	1.13	0.79	0.68
Intercept [L/kg_{VS}]	-9.72	-122	138	167
Bias [L/kg_{VS}]	-11.1	-39.9	9.11	-3.81
r	0.85	0.75	0.65	0.62
R^2	0.72	0.57	0.43	0.38
RMSE [L/kg_{VS}]	60.6	74.8	72.9	59.8
CVRMSE [%]	9.77	11.9	12.0	10.4

The validation was performed for each individual sample. Based on all four validations, it can be concluded that the model depicted the variation of the datasets with an estimation error (CVRMSE) of about 11 % and with a moderate correlation between the measured and the predicted values.

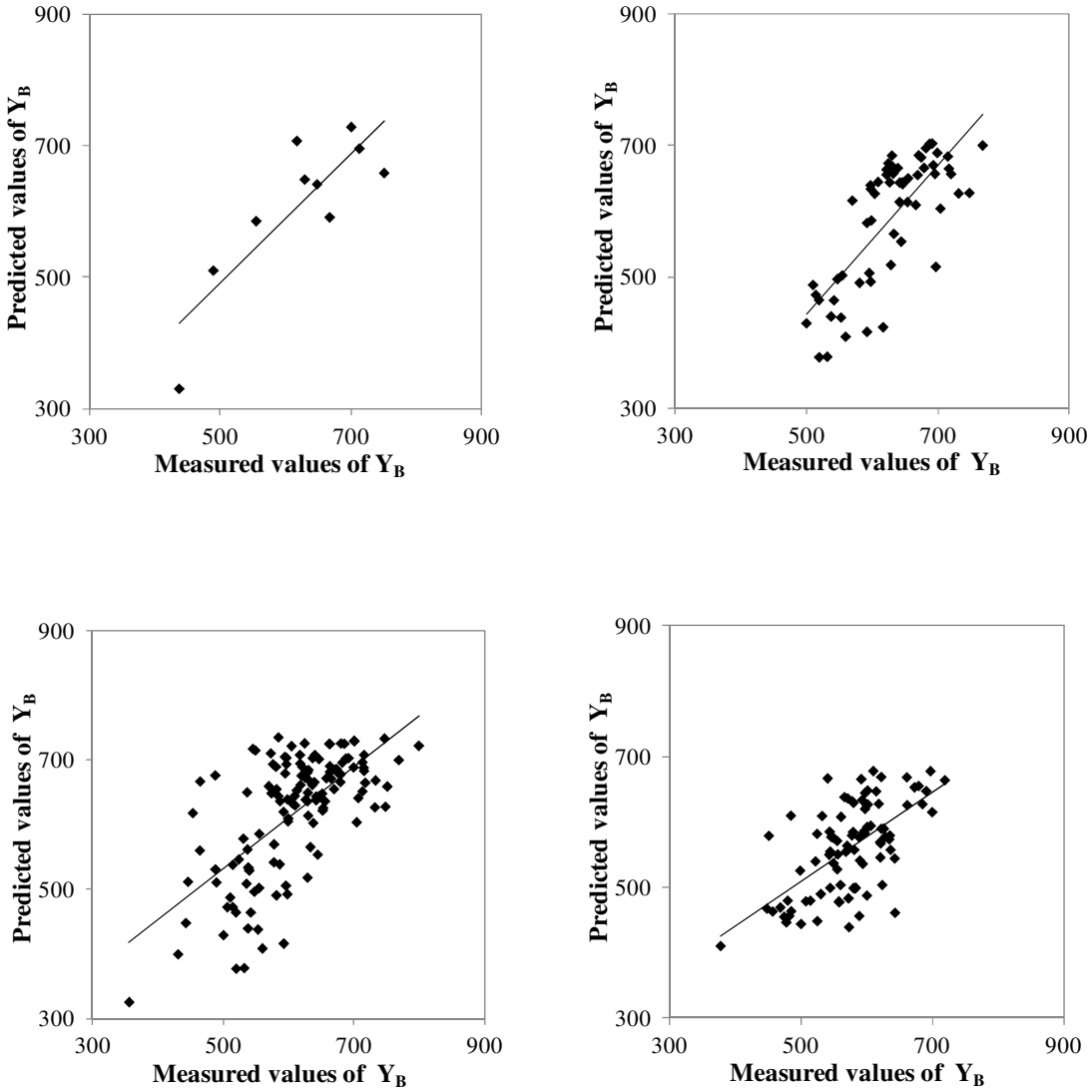

Figure 5: Performance of the across plant species model tested with different independent datasets; measured versus predicted values of biogas yield. Top left: first dataset, top right: second dataset, lower left: third dataset and lower right: forth dataset.

Based on the validation with the first dataset, the regression line indicated a precise prediction with the values for the slope, intercept and bias to be very close to the optimum values. The correlation between the measured and predicted values was strong ($r = 0.85$) and the R^2 indicated that 72 % of the total validation of the dataset was

explained. With this dataset the best values of external validation were achieved (Table 11). Although the dataset was small and similar to the calibration's dataset, the validation revealed the high potential of the model.

Based on the validation with the second dataset, the regression line between the measured and predicted values indicated a slight underestimation (bias = - 39.9) of the prediction. However, since the slope is above 1 and the intercept below 0, the model underestimated the measured values below the average and overestimated the measured values above the average. The correlation between the measured and predicted values was moderate (r = 0.75) and the R^2 was 0.57 (Figure 5, Table 11).

For the third and fourth validation datasets, similar statistical values were observed. Both datasets consisted of samples of various plant species and plant groups. The bias value was very small and the estimation error was low (CVRMSE of 12.0 % and 10.4 %, respectively). This indicates that the average prediction of the dataset was precise. However, the prediction of each individual sample was not sufficient with R^2 values below 0.5.

In summary, predicting the behavior of a biological system is very challenging. The model can predict the variation of various datasets with an estimation error of 11 %, and it can expose the differentiation of the samples on the biogas yield within a dataset. However, the biogas yield prediction of each single sample within a dataset was not sufficient.

7.1.2 Grassland plant species model

The specific model for grassland plant species was validated with a cross-validation method, as detailed in Chapter 6.2. The results were satisfying; however, an external validation would be desirable to confirm the results. In Chapter 6.4, an independent dataset of 55 grassland samples was analyzed to validate the grassland species model. The plant species selected for the model validation were the same plant species as were used for calibration, but they were cultivated and collected in a different year (2014).

In total, four models were assessed and validated. The across plant species model (Dandikas et al., 2014), the grassland species specific model (Dandikas et al., 2015), and two existing models developed for energy crops (global models) were selected to compare the results of the validation as developed by Triolo et al. (2011) and Thomsen et al. (2014). Both models were created to predict only the methane yield and thus not the biogas yield. Therefore, only the values of the methane yield were presented during the validation.

The validation of the models was performed with two different approaches. The first approach was a validation of each individual sample (n = 55), and the second approach with the mean value of each cultivar (n = 9). During the first approach the effect of each harvest date and cut were emphasized; during the second approach the differences of the cultivars were emphasized. According to the BMP tests and the fodder analysis, significant differences between the datasets from 2013 and 2014 were observed. The main reason for this effect was most likely the different weather conditions in 2013 and 2014 (Figure 2).

During the single sample validation approach, none of the models could explain more than 27 % of the dataset variation, despite the estimation error (CVRMSE) being between 8 and 13%. The correlation between the measured and the predicted values can be characterized as weak, since the r value varied between 0.39 and 0.52. Similar results were published by Rath et al. (2015). An explanation for this observation could be that each value in the dataset was an individual sample, leading to a high uncertainty. Possible reasons for this are: Firstly, the sample could not be considered representative due to the microclimate conditions and/or other external effects that caused the chemical structure of the sample to be atypical for this specific harvest date or phenological stage (defined BBCH code). In both years (2013 and 2014), each harvest date was defined at a specific BBCH code; however, due to different weather conditions the phenological stages could have been shifted or not visually well recognized. Secondly, the precision and accuracy of the chemical analysis (fodder analysis) and moreover the batch test (biological test) can be affected by many factors. These two reasons caused model uncertainty and this could be minimized by the use of average values with respect to the plant cultivars.

The methane yield prediction performance of each model for each individual sample can be summarized as follows:

Model of Triolo et al. (2011): The average value of the predicted methane yield of the dataset was very close to the measured methane yield. Although the maximum value was overestimated and the minimum value was underestimated, which led to a higher range of Y_M, the slope of the regression line was 0.71 and was the best value among the three following models. Since the values were close to the best fit line, the estimation error (CVRMSE) was 11 %.

Model of Thomsen et al. (2014): The average value of the predicted methane yield of the dataset was very close to the measured methane yield, as well. However, the maximum value was underestimated and the minimum value was overestimated. This led to an impressive reduced range of Y_M values. The slope of the regression line was 0.19, and the variation among the samples could not be shown. However, since the values were close to the best fit line, the estimation error (CVRMSE) reached only 8 %, which is the best value compared to the values of the other models.

Model of Dandikas et al. (2014): The average value of the predicted methane yield of the dataset was very close to the measured methane yield. The maximum value of the dataset was slightly overestimated and the minimum value was underestimated; however, the range of the values could be predicted with only an 18 L/kg_{VS} difference, thus the variation of the dataset could be well predicted. The slope of the regression line was 0.47 and the estimation error (CVRMSE) was 11 %.

Model of Dandikas et al. (2015): The correlation between the measured and predicted values was 0.44 and the values of the methane yield were always overestimated with a bias of 32 L/kg_{VS}. The slope of the regression line was 0.34 and the estimation error (CVRMSE) was 13 %, since the values were systematically above the perfect fit line.

For the second validation approach, the average values of each cultivar were used. All four models showed a better performance using average values; however, the results revealed differences among the models. A high correlation ($r > 0.85$) between the predicted and the measured values was observed for all four models, and thus more than 70 % of the dataset variation could be explained ($R^2 > 0.7$). The global models of Triolo at al. (2011) and Dandikas et al. (2014) showed similar behavior with similar values for the statistical parameters (Table 4, in **Paper IV**). Both models could expose the two plant groups, grass and legume species, and both models performed better for the grass species than for the legume species (their values were underestimated). The model of Thomsen at al. 2014 could also expose the difference between the two plant groups, but the range of the values was underestimated. Within the two plant groups similar values were predicted and almost no difference could be seen. The regression line of the model of Dandikas et al. (2015), which was a grassland species specific model, showed the best performance with the predicted values significantly correlated with the measured values. However, the values were systematically overestimated, probably due to external factors, and a bias correction was applied.

The prediction performance of each model of the methane yield from the nine cultivars can be summarized as follows:

Model of Triolo et al. (2011): The predicted values were significantly correlated with the measured values ($r = 0.90$). The average value of the predicted methane yield of the dataset was very close to the measured average methane yield. However, the slope of the regression line was 1.80. This means the maximum value was overestimated, and the minimum value was underestimated. The variation among the samples could be well predicted, but the range of the values was overestimated. The estimation error (CVRMSE) was 6 %.

Model of Thomsen et al. (2014): This model also showed similar performance with the average values of the samples. The correlation between the predicted and the measured values was 0.89, but the slope of the regression line was 0.52. Therefore, the maximum

value was underestimated and the minimum value was overestimated, leading to a lower range of the values. However, since the values were close to the best fit line, the estimation error (CVRMSE) was 3 %.

Model of Dandikas et al. (2014): The predicted values were significantly correlated with the measured values (r = 0.85). The average value of the predicted methane yield of the dataset was very close to the measured average methane yield. However, the slope of the regression line was 1.69. This means the maximum value was overestimated and the minimum value was underestimated. The variation among the samples could be well predicted, but the range of the values was overestimated. The estimation error (CVRMSE) was 6 %.

Model of Dandikas et al. (2015): The predicted values were significantly correlated with the measured values with the best r value of 0.92. Moreover, the slope of the regression line was 1, which means all values were predicted accurately. However, the whole dataset was systematically overestimated with a bias of 33 L/kg_{VS}. Since the values were systematically deviated from the best fit line, the estimation error (CVRMSE) was 10 %. An overestimation of the values was to be expected, as the values of the dataset for the calibration were higher than the values of the dataset for the validation. This can be characterized as systematic error, and therefore, a bias correction can be conducted. The bias correction reduced the estimation error to 2 %.

Finally, based on the external validation, it could be concluded that: Firstly, regression models can reveal the plant differentiation on methane yield; secondly, average values are needed for estimation error reduction, since average values reduce the analytical error and increase the accuracy of the model; and thirdly, external parameters (e.g. weather conditions) have a significant influence on the predicted values.

7.1.3 Hydrolysis rate constant model

The model developed for predicting the hydrolysis rate constant was introduced in **Paper III**. The average prediction error (CVRMSE) was 15 % for an external dataset. The range of the dataset was underestimated, and the CV was 10 points lower than the one of the measured values. However, the mean of the predicted values was similar to the mean of the measured values, and the predicted values were strongly correlated with the measured values as indicated by a correlation coefficient of 0.93. The results indicate that the $k_{h\_0.5}$ model could predict the variation in the dataset very well. This is further manifested in the plot of measured versus predicted values, as the points were very close to their regression line with a R^2 value of 0.86 (Figure 3 in **Paper III**). The model has proven its suitability as a global model since it was able to predict the BMP curve of unknown species (not included in calibration dataset). Although the values

could not be estimated with high accuracy, feedstock ranking according to their BMP and hydrolysis rate constant was possible. The biochemical processes of anaerobic digestion are very complex and cannot be precisely described by a simple first-order kinetic model, as has already been concluded by Li et al. (2016). However, the models could accurately predict the differentiation of the samples and these results depict the main advantage of this study's approach.

7.2 Chemical compounds as regressors

Fodder analysis is a well-established standard method for animal feed assessment. Round robin tests have shown that although the relative standard deviation for the chemical analyses among laboratories was high (Henkelmann and Fischer-Kaiser, 2014), the deviation within the laboratory was low and the repeatability lain within the acceptable range. This fact makes the analytical parameters suitable for statistical analysis and several authors have reported a significant correlation between the chemical compounds and the biogas yield. However statistically, the properties of the chemical compound cannot be considered related to the methane yield, but rather the proportion of the chemical compound included. Moreover, it needs to be considered that any change in the content of one compound will change the content of the others, due to the fact that they are recorded in percentage of total solids.

In this study, it has been shown that the biogas yield can be predicted by ADL and HC for a global model, and XP needs to be included for grassland samples to reduce the prediction error. For the prediction of the hydrolysis rate constant, XP and NFC are needed. Each of these regressors contributed differently to the MLR model. Below, the importance of each analytical parameter for the prediction of the biogas yield is discussed.

ADL has a strong negative effect on the biogas yield. The standardized regression coefficient of the models confirmed that ADL is the most important regressor. Several authors have confirmed the negative correlation of ADL with the biogas yield (Gunaseelan, 2009; Rath et al., 2013). Other authors did not use ADL as a regressor (Amon et al., 2007b; Weißbach, 2009b). However, this does not exclude, at least indirectly, the negative effect of ADL on the biogas yield. The analytical parameter ADL represents the complex lignocellulosic structure, which means lignin as well as a portion of HC and CL that is hardly bioavailable. Thus, the regressor ADL reflects the negative influence of the cell-wall components on the biogas yield.

In this study, HC was the second most important regressor, and it was used in MLR to explain the positive effect of the fiber matrix on biogas production. Hemicellulose and cellulose are structured together to be part of the cell-wall. However, hemicellulose is

available faster, since it can be faster diluted in water, and it was expected to be a readily digestible fiber fraction. The analytical parameter HC is defined as the difference between NDF and ADF. Thomsen et al. (2014) also reported the positive effect of HC on Y_B.

XP was utilized in MLR to improve the prediction accuracy of the grassland species specific model. Despite the low potential contribution on the biogas yield of XP (since the absolute amount of XP is low), XP depicts the variation in Y_B within the grassland species. XP was also statistically significant for the prediction model of $k_{h\_0.5}$. Although the hydrolysis rate of proteins is known to be lower than for carbohydrates (Lübken et al., 2015), the analytical parameter XP characterizes the ageing of the crop, with young plants characterized by a high share of XP and a good anaerobic digestibility and vice versa.

NFC also has a positive effect on the prediction of $k_{h\_0.5}$. Carbohydrates (fiber and non-fiber) are the main feed source for biogas production, since they occupy more than 80 % of the total VS of energy crops. During the anaerobic process in a BMP test, the readily digestible fractions are degraded first, and these are mostly characterized by NFC.

7.3 Power and limitation of BMP prediction

As it was discussed previously, a BMP test is a costly and time-consuming examination method and this disadvantage led to the need for a new method to obtain the same information. Strömberg et al. (2015) suggested reducing the experimental duration of a BMP test and modeling the ultimate methane production. However, the uncertainty of the prediction still exists and a proper experimental set up is still needed. The BMP prediction based on the chemical composition of the crops is a fast and low-cost method. But, these regression models also have disadvantages. Therefore, the models have to be studied with external validations, in order to define their capabilities and limits. The uncertainty of the analytical parameters of the chemical compounds also has to be considered. At this point it should be clear that a chemical analysis can be defined quite well when the measurement can be repeated. But the biological analytical method of a BMP test cannot be expected to be exactly the same for each assay. Moreover, the effect of the environmental factors on the crops can be reflected by their chemical compositions, since these vary with respect to plant species, plant maturity and climate conditions (Wahid et al., 2015). This cause of variation must be regularly included into the calibration dataset.

The BMP prediction based on the chemical compounds of the feedstock is also challenging, since co-effects among the chemical compounds always exist. Bekiaris et al. (2015) showed the importance of lignin and hemicelluloses in the

protection of cellulose in respect to its bioavailability. Mussatto et al. (2008) observed an increase in the cellulose conversion into glucose when hemicelluloses and lignin were removed. They concluded that the lower the hemicellulose and lignin contents are in the sample, the higher the efficiency of cellulose hydrolysis will be. Hence, the cellulose hydrolysis was affected by the presence of hemicellulose and lignin in the sample. The fiber matrix (lignocellulosic biomass) influences the velocity and the degree of degradation, and even more the specific biogas production. However, the fiber matrix is a plant species specific trait. This indicates the necessity for plant species specific models; however, this is an expensive procedure in terms of time and resources.

The first-order kinetic model was used for the prediction of the hourly biogas production. In this study the suitability of the first-order model has been shown. However, it can be found in literature that the difference between the predicted and measured biogas yields (fitting error) was higher with the first-order kinetic model than with the modified Gompertz model for non-fiber biomass (Yono et al., 2014) and for fiber biomass (Tsapekos et al., 2017). For agricultural residuals, it has been also reported that the conventional first-order hydrolysis kinetic model was not suitable for describing the entire hydrolysis process of maize stover, because there were two first-order decay periods for hydrolysis (Li et al., 2016). Although, the estimation error of this approach is high, it is suitable for a global prediction model.

8 Outlook

The simulation of the biogas production process as a white box is extremely demanding (Mulka et al., 2016) due to the high variability of microbes, the feedstock specific microbial communities involved, and the different optimal growth conditions (Batstone et al., 2015; Blasco et al., 2014; Lebuhn et al., 2014). The anaerobic digestion model no. 1 (ADM1) describes biochemical and physicochemical processes resulting in a large number of stoichiometric and kinetic equations (Batstone et al., 2002). The complexity of ADM1 leads to the need for many input parameters and often their identification is difficult or not possible (Donoso-Bravo et al., 2011). However, considering the anaerobic process as a black box is a much more simple approach and allows the prediction of the biogas production. As these models disregard reaction mechanisms, they would be more suitable to control AD processes, rather than to design and scale them up (Yu et al., 2013). Simple calculators mainly use the relation that exists between volatile solids and biogas production. Nevertheless, the aim of these calculators is not to simulate the process of anaerobic digestion, but to estimate the applicability of the process to a specific farm and to provide information to a farmer or a decision maker (Kythreotou et al., 2014).

In this study, the biochemical anaerobic process was considered as a black box system and the BMP test was explained in terms of input (feedstock) and output (biogas). Moreover, a simple first-order model was used to estimate the kinetics of a BMP test. It has been proven that for various energy crops the proposed approach reflected the differentiation of the samples. A large dataset for the calibration and the validation of the model was used. Hence, the model should be considered as a global model. Future studies could enhance the existing model by extending the calibration dataset with additional plant species or cultivars. Specific models can be developed following the proposed approach in order to identify variables that could minimize the estimation error. New analytical parameters can also be included, such as a non-fodder analysis parameter like the high calorific value (HCV) for BMP prediction (Edwiges et al., 2018). In addition, the identification of rapidly digestible and slowly digestible fractions is crucial for future research (Weinrich and Nelles, 2015), as also supported by Li et al. (2016) having also proposed two first-order decay periods for hydrolysis.

9 References

Amon, T., 2007. Methanertragspotential verschiedener Energiepflanzen – Aktueller Stand. Energiepflanzenkolloquium KWS SAAT AG.

Amon, T., Amon, B., Kryvoruchko, V., Machmüller, A., Hopfner-Sixt, K., Bodiroza, V., Hrbek, R., Friedel, J., Pötsch, E., Wagentristl, H., Schreiner, M., Zollitsch, W., 2007a. Methane production through anaerobic digestion of various energy crops grown in sustainable crop rotations. Bioresour. Technol. 98, 3204–3212.

Amon, T., Amon, B., Kryvoruchko, V., Zollitsch, W., Mayer, K., Gruber, L., 2007b. Biogas production from maize and dairy cattle manure – Influence of biomass composition on the methane yield. Agric. Ecosyst. Environ. 118, 173–182.

Angelidaki, I., Alves, M., Bolzonella, D., Borzacconi, L., Campos, J.L., Guwy, A.J., Kalyuzhnyi, S., Jenicek, P., van Lier, J.B., 2009. Defining the biomethane potential (BMP) of solid organic wastes and energy crops: a proposed protocol for batch assays. Water Sci. Technol. 59, 927–934.

Baserga, U., 1998. Landwirtschaftliche Co-Vergärungs-Biogasanlagen. Biogas aus organischen Reststoffen und Energiegras. Agricultural co-digestion biogas plants. Biogas from organic waste and energy grass. FAT-Berichte Nr 512.

Batstone, D.J., Keller, J., Angelidaki, I., Kalyuzhnyi, S.V., Pavlostathis, S.G., Rozzi, A., Sanders, W.T.M., Siegrist, H., Vavilin, V.A., 2002. The IWA Anaerobic Digestion Model No 1 (ADM1). Water Sci. Technol. 45, 65–73.

Batstone, D.J., Puyol, D., Flores-Alsina, X., Rodríguez, J., 2015. Mathematical modelling of anaerobic digestion processes: applications and future needs. Rev. Environ. Sci. Biotechnol. 14, 595–613.

Bauer, C., Lebuhn, M., Gronauer, A., 2009. Mikrobiologische Prozesse in landwirtschaftlichen Biogasanlagen. LfL-Schriftenreihe.

Bekiaris, G., Triolo, J.M., Peltre, C., Pedersen, L., Jensen, L.S., Bruun, S., 2015. Rapid estimation of the biochemical methane potential of plant biomasses using Fourier transform mid-infrared photoacoustic spectroscopy. Bioresour. Technol. 197, 475–481.

Bere, R.C., Precup, I.B., Silvestru, C.I., 2015. On Growth Poles from EU Countries in the Framework of Europe 2020. Procedia Econ. Finance, 2nd Global Conference on Business, Economics, Management and Tourism 23, 920–925.

Blasco, L., Kahala, M., Tampio, E., Ervasti, S., Paavola, T., Rintala, J., Joutsjoki, V., 2014. Dynamics of microbial communities in untreated and autoclaved food waste anaerobic digesters. Anaerobe 29, 3–9.

Boyle, W.C., 1976. Energy recovery from sanitary landfills – a review. Schlegel H G Barnea Hrsg Microb. Energy Convers. Oxf. Pergamon Press 119–138.

Buswell, A.M., Mueller, H.F., 1952. Mechanism of Methane Fermentation. Ind. Eng. Chem. 44, 550–552.

Chandra, R., Takeuchi, H., Hasegawa, T., 2012. Methane production from lignocellulosic agricultural crop wastes: A review in context to second generation of biofuel production. Renew. Sustain. Energy Rev. 16, 1462–1476.

Dandikas, V., Heuwinkel, H., Lichti, F., Drewes, J.E., Koch, K., 2014. Correlation between biogas yield and chemical composition of energy crops. Bioresource Technology 174, 316–320.

Dandikas, V., Heuwinkel, H., Lichti, F., Drewes, J.E., Koch, K., 2015. Correlation between Biogas Yield and Chemical Composition of Grassland Plant Species. Energy Fuels 29, 7221–7229.

Dandikas, V., Heuwinkel, H., Lichti, F., Drewes, J.E., Koch, K., 2018a. Predicting methane yield by linear regression models: A validation study for grassland biomass. Bioresource Technology 265, 372–379.

Dandikas, V., Heuwinkel, H., Lichti, F., Eckl, T., Drewes, J.E., Koch, K., 2018b. Correlation between hydrolysis rate constant and chemical composition of energy crops. Renewable Energy 118, 34–42.

Dollhofer, V., Callaghan, T.M., Griffith, G.W., Lebuhn, M., Bauer, J., 2017. Presence and transcriptional activity of anaerobic fungi in agricultural biogas plants. Bioresour. Technol. 235, 131–139.

Donoso-Bravo, A., Mailier, J., Martin, C., Rodríguez, J., Aceves-Lara, C.A., Wouwer, A.V., 2011. Model selection, identification and validation in anaerobic digestion: A review. Water Res. 45, 5347–5364.

Edwiges, T., Frare, L., Mayer, B., Lins, L., Mi Triolo, J., Flotats, X., de Mendonça Costa, M.S.S., 2018. Influence of chemical composition on biochemical methane potential of fruit and vegetable waste. Waste Manag. 71, 618–625.

Ellabban, O., Abu-Rub, H., Blaabjerg, F., 2014. Renewable energy resources: Current status, future prospects and their enabling technology. Renew. Sustain. Energy Rev. 39, 748–764.

Esen, M., Yuksel, T., 2013. Experimental evaluation of using various renewable energy sources for heating a greenhouse. Energy Build. 65, 340–351.

Gerardi, M.H., 2003. The microbiology of anaerobic digesters. Wiley-IEEE.

Gu, Y., Chen, X., Liu, Z., Zhou, X., Zhang, Y., 2014. Effect of inoculum sources on the anaerobic digestion of rice straw. Bioresour. Technol. 158, 149–155.

Gunaseelan, V.N., 2009. Predicting ultimate methane yields of Jatropha curcus and Morus indica from their chemical composition. Bioresour. Technol. 100, 3426–3429.

Henkelmann, G., Fischer-Kaiser, K., 2014. Die Qualität von Faseranalysen und Untersuchungen im Umfeld von Biogasanlagen, in: Kongressband – Untersuchen, Bewerten, Beraten, Forschen: 125 Jahre VDLUFA Im Dienste von Landwirtschaft , Umwelt- Und Verbraucherschutz. VDLUFA-Verlag, Darmstadt, pp. 845–853.

Hess, M., Barralis, G., Bleiholder, H., Buhr, L., Eggers, T., Hack, H., Stauss, R., 1997. Use of the extended BBCH scale – general for the descriptions of the growth stages of mono; and dicotyledonous weed species. Weed Res. 37, 433–441.

Heuwinkel, H., Aschmann, A., Gerlach, R., Gronauer, A., 2009. Die Genauigkeit der Messung des Gasertragspotentials von Substraten mit der Batchmethode. In proceedings of Biogas Science, Erding, Germany, 2009 Vol. 1, 95–103.

Holliger, C., Alves, M., Andrade, D., Angelidaki, I., Astals, S., Baier, U., Bougrier, C., Buffière, P., Carballa, M., Wilde, V. de, Ebertseder, F., Fernández, B., Ficara, E., Fotidis, I., Frigon, J.-C., Laclos, H.F. de, Ghasimi, D.S.M., Hack, G., Hartel, M., Heerenklage, J., Horvath, I.S., Jenicek, P., Koch, K., Krautwald, J., Lizasoain, J., Liu, J., Mosberger, L., Nistor, M., Oechsner, H., Oliveira, J.V., Paterson, M., Pauss, A., Pommier, S., Porqueddu, I., Raposo, F., Ribeiro, T., Pfund, F.R., Strömberg, S., Torrijos, M., Eekert, M. van, Lier, J. van,

Wedwitschka, H., Wierinck, I., 2016. Towards a standardization of biomethane potential tests. Water Sci. Technol. 74, 2515–2522.

Jolliffe, I.T., 2002. Principal Component Analysis. Springer, New York.

Jung, H.-J.G., 1997. Analysis of Forage Fiber and Cell Walls in Ruminant Nutrition. J. Nutr. 127, 810–813.

Kaiser, F.L., 2007. Einfluss der stofflichen Zusammensetzung auf die Verdaulichkeit nachwachsender Rohstoffe beim anaeroben Abbau in Biogasreaktoren., PhD Thesis. TU München, München.

Karakashev, D., Batstone, D.J., Angelidaki, I., 2005. Influence of Environmental Conditions on Methanogenic Compositions in Anaerobic Biogas Reactors. Appl. Environ. Microbiol. 71, 331–338.

Keymer, U., Schilcher, A., 1999. Überlegungen zur Errechnung theoretischer Gasausbeuten vergärbarer Substrate in Biogasanlagen. Considerations for calculating the theoretical gas yield of digestible substrates for biogas plants. Landtechik-Bericht Nr. 32.

Koch, K., Drewes, J.E., 2014. Alternative approach to estimate the hydrolysis rate constant of particulate material from batch data. Appl. Energy 120, 11–15.

Kortsch, T., Hildebrand, J., Schweizer-Ries, P., 2015. Acceptance of biomass plants – Results of a longitudinal study in the bioenergy-region Altmark. Renew. Energy 83, 690–697.

Kythreotou, N., Florides, G., Tassou, S.A., 2014. A review of simple to scientific models for anaerobic digestion. Renew. Energy 71, 701–714.

Lebuhn, M., Hanreich, A., Klocke, M., Schlüter, A., Bauer, C., Pérez, C.M., 2014. Towards molecular biomarkers for biogas production from lignocellulose-rich substrates. Anaerobe 29, 10–21.

Li, D., Huang, X., Wang, Q., Yuan, Y., Yan, Z., Li, Z., Huang, Y., Liu, X., 2016. Kinetics of methane production and hydrolysis in anaerobic digestion of corn stover. Energy 102, 1–9.

Li, Y., Zhang, R., Liu, G., Chen, C., He, Y., Liu, X., 2013. Comparison of methane production potential, biodegradability, and kinetics of different organic substrates. Bioresour. Technol. 149, 565–569.

Liobikienė, G., Butkus, M., 2017. The European Union possibilities to achieve targets of Europe 2020 and Paris agreement climate policy. Renew. Energy 106, 298–309.

Lopes, W.S., Leite, V.D., Prasad, S., 2004. Influence of inoculum on performance of anaerobic reactors for treating municipal solid waste. Bioresour. Technol. 94, 261–266.

Lübken, M., Gehring, T., Wichern, M., 2010. Microbiological fermentation of lignocellulosic biomass: current state and prospects of mathematical modeling. Appl. Microbiol. Biotechnol. 85, 1643–1652.

Lübken, M., Koch, K., Gehring, T., Horn, H., Wichern, M., 2015. Parameter estimation and long-term process simulation of a biogas reactor operated under trace elements limitation. Appl. Energy 142, 352–360.

Meier, U. 2018, Growth stages of mono- and dicotyledonous plants: BBCH Monograph, Julius Kühn-Institut (JKI), Quedlinburg, Germany.

Meyer, A.K.P., Ehimen, E.A., Holm-Nielsen, J.B., 2018. Future European biogas: Animal manure, straw and grass potentials for a sustainable European biogas production. Biomass Bioenergy. 111, 154–164.

Meyer-Aurich, A., Schattauer, A., Hellebrand, H.J., Klauss, H., Plöchl, M., Berg, W., 2012. Impact of uncertainties on greenhouse gas mitigation potential of biogas production from agricultural resources. Renew. Energy 37, 277–284.

Mittweg, G., Oechsner, H., Hahn, V., Lemmer, A., Reinhardt-Hanisch, A., 2012. Repeatability of a laboratory batch method to determine the specific biogas and methane yields. Eng. Life Sci. 12, 270–278.

Mulka, R., Szulczewski, W., Szlachta, J., Mulka, M., 2016. Estimation of methane production for batch technology – A new approach. Renew. Energy 90, 440–449.

Munk, B., Bauer, C., Gronauer, A., Lebuhn, M., 2010. Population dynamics of methanogens during acidification of biogas fermenters fed with maize silage. Eng. Life Sci. 10, 496–508.

Mussatto, S.I., Fernandes, M., Milagres, A.M.F., Roberto, I.C., 2008. Effect of hemicellulose and lignin on enzymatic hydrolysis of cellulose from brewer's spent grain. Enzyme Microb. Technol. 43, 124–129.

Pavlostathis, S.G., Giraldo-Gomez, E., 1991. Kinetics of anaerobic treatment: A critical review. Crit. Rev. Environ. Control 21, 411–490.

Pobeheim, H., Munk, B., Lindorfer, H., Guebitz, G.M., 2011. Impact of nickel and cobalt on biogas production and process stability during semi-continuous anaerobic fermentation of a model substrate for maize silage. Water Res. 45, 781–787.

Raposo, F., Fernández-Cegrí, V., De la Rubia, M. a., Borja, R., Béline, F., Cavinato, C., Demirer, G., Fernández, B., Fernández-Polanco, M., Frigon, J. c., Ganesh, R., Kaparaju, P., Koubova, J., Méndez, R., Menin, G., Peene, A., Scherer, P., Torrijos, M., Uellendahl, H., Wierinck, I., de Wilde, V., 2011. Biochemical methane potential (BMP) of solid organic substrates: evaluation of anaerobic biodegradability using data from an international interlaboratory study. J. Chem. Technol. Biotechnol. 86, 1088–1098.

Rath, J., Heuwinkel, H., Herrmann, A., 2013. Specific Biogas Yield of Maize Can Be Predicted by the Interaction of Four Biochemical Constituents. BioEnergy Res. 6, 939–952.

Rath, J., Heuwinkel, H., Taube, F., Herrmann, A., 2015. Predicting Specific Biogas Yield of Maize-Validation of Different Model Approaches. BioEnergy Res. 8, 832–842.

Strömberg, S., Nistor, M., Liu, J., 2015. Early prediction of Biochemical Methane Potential through statistical and kinetic modelling of initial gas production. Bioresour. Technol. 176, 233–241.

Tabachnick, B.G., Fidell, L.S., 2012. Using Multivariate Statistics. Pearson Education, Boston.

Thomsen, S.T., Spliid, H., Østergård, H., 2014. Statistical prediction of biomethane potentials based on the composition of lignocellulosic biomass. Bioresour. Technol. 154, 80–86.

Triolo, J.M., Sommer, S.G., Møller, H.B., Weisbjerg, M.R., Jiang, X.Y., 2011. A new algorithm to characterize biodegradability of biomass during anaerobic digestion: Influence of lignin concentration on methane production potential. Bioresour. Technol. 102, 9395–9402.

Tsapekos, P., Kougias, P.G., Egelund, H., Larsen, U., Pedersen, J., Trénel, P., Angelidaki, I., 2017. Mechanical pretreatment at harvesting increases the bioenergy output from marginal land grasses. Renew. Energy 111, 914–921.

Uusitalo, V., Soukka, R., Horttanainen, M., Niskanen, A., Havukainen, J., 2013. Economics and greenhouse gas balance of biogas use systems in the Finnish transportation sector. Renew. Energy 51, 132–140.

VDI, 2006. Guideline VDI 4630 – Fermentation of organic materials Characterisation of the substrate, sampling, collection of material data, fermentation tests. The Association of German Engineers, Düsseldorf, Germany.

VDI, 2016. Guideline VDI 4630 – Fermentation of organic materials Characterisation of the substrate, sampling, collection of material data, fermentation tests. The Association of German Engineers, Düsseldorf, Germany.

VDLUFA, 1976. Die chemische Untersuchung von Futtermitteln. (engl.: Chemical Analysis of Feedstuffs.), Band 3. VDLUFA, Darmstadt, Germany.

VDLUFA, 2011. Bestimmung der Biogas- und Methanausbeute in Gärtests. (engl.: Measurement of Biogas and Methane Yields in Fermentation Tests). Band 7 - 4.1.1. VDLUFA, Darmstadt, Germany.

Wahid, R., Ward, A.J., Møller, H.B., Søegaard, K., Eriksen, J., 2015. Biogas potential from forbs and grass-clover mixture with the application of near infrared spectroscopy. Bioresour. Technol. 198, 124–132.

Weinrich, S., Nelles, M., 2015. Critical comparison of different model structures for the applied simulation of the anaerobic digestion of agricultural energy crops. Bioresour. Technol. 178, 306–312.

Weißbach, F., 2008. Zur Bewertung des Gasbildungspotenzials von nachwachsenden Rohstoffen. Landtechnik 63, 356–358.

Weißbach, F., 2009a. Die Bewertung des Gasbildungspotenzials von nachwachsenden Rohstoffen. In proceedings of Biogas Science, Erding, Germany, 2009, Vol. 3, 517–526.

Weißbach, F., 2009b. Evaluation of the renewable primary products for biogas production. Part I: Gas production potential of the fermentable nutrients. Pflanzenbauwissenschaften 13, 72–85.

Xu, F., Wang, Z.-W., Li, Y., 2014. Predicting the methane yield of lignocellulosic biomass in mesophilic solid-state anaerobic digestion based on feedstock characteristics and process parameters. Bioresour. Technol. 173, 168–176.

Yono, B., Syaichurrozi, I., Sumardiono, S., 2014. Kinetic Model of Biogas Yield Production from Vinasse at Various Initial pH: Comparison between Modified Gompertz Model and First Order Kinetic Model. Res. J. Appl. Sci. Eng. Technol. 7, 2798–2805.

Yu, L., Wensel, P.C., Ma, J., Chen, S., 2013. Mathematical Modeling in Anaerobic Digestion (AD). J. Bioremediation Biodegrad. 0, 1–12.

Zervos, A., Lins, C., Tesniere, L., 2011. Mapping Renewable Energy Pathways towards 2020 - EU Roadmap. European Renewable Energy Council, ETM Library, Brussels.

10 Appendix

- Paper I
- Paper II
- Paper III
- Paper IV

Paper I

Bioresource Technology 174 (2014) 316–320

Contents lists available at ScienceDirect

Bioresource Technology

journal homepage: www.elsevier.com/locate/biortech

Short Communication

Correlation between biogas yield and chemical composition of energy crops

V. Dandikas [a,*], H. Heuwinkel [b], F. Lichti [a], J.E. Drewes [c], K. Koch [c]

[a] Institute for Agricultural Engineering and Animal Husbandry, Bavarian State Research Center for Agriculture, Am Staudengarten 3, 85354 Freising, Germany
[b] Department of Agriculture and Food Economy, Hochschule Weihenstephan-Triesdorf, Am Hofgarten 1, 85354 Freising, Germany
[c] Chair of Urban Water Systems Engineering, Technische Universität München, Am Coulombwall 8, 85748 Garching, Germany

HIGHLIGHTS

- 41 energy crops were analyzed in batch anaerobic digestion tests.
- Furthermore, the chemical composition of all samples was determined.
- 80% of the sample variation on biogas yield can be explained through lignin.
- Lignin and hemicellulose are suggested as model variables.

ARTICLE INFO

Article history:
Received 4 August 2014
Received in revised form 1 October 2014
Accepted 5 October 2014
Available online 13 October 2014

Keywords:
Batch test
Biogas yield potential
Empirical model
Lignocellulose
Regression analysis

ABSTRACT

The scope of this study was to investigate the influence of the chemical composition of energy crops on biogas and methane yield. In total, 41 different plants were analyzed in batch test and their chemical composition was determined. For acid detergent lignin (ADL) content below 10% of total solids, a significant negative correlation for biogas and methane yields ($r \approx -0.90$) was observed. Based on a simple regression analysis, more than 80% of the sample variation can be explained through ADL. Based on a principal component analysis and multiple regression analysis, ADL and hemicellulose are suggested as suitable model variables for biogas yield potential predictions across plant species.

1. Introduction

In order to maintain a high cost-efficiency of agricultural biogas plants, a flexible operation using different feed stocks is desired. Previous studies have reported that long-term mono-digestion of energy crops can result in biochemical or mechanical process instability (De Moor et al., 2013; Koch et al., 2009; Lebuhn et al., 2008). Anaerobic degradation is a process where microbiological and chemical aspects are closely linked (Angelidaki et al., 2009; Mittweg et al., 2012). Additionally, the chemical composition of the plant species and, more specifically, the lignocellulosic matrix can affect the potential biogas yield (Li et al., 2013; Lübken et al., 2010). Determining the potential biogas yield of different substrates *a priori* can guide a more efficient operation of biogas facilities. Anaerobic batch tests under standardized laboratory conditions offer information about the potential biogas and methane yield (Angelidaki et al., 2009; Koch and Drewes, 2014). However, conducting batch tests is complex and time consuming.

To estimate quickly and reliably the biogas and methane yield potential, empirical models were developed based on the chemical composition of the biomass and experimental values of biogas yield (Y_B). Amon et al. (2007) developed a model based on maize samples, where crude protein, crude lipids, cellulose and hemicellulose had a positive impact on methane yield (Y_M). For the differentiation of maize cultivars, Rath et al. (2013) found a negative influence of lignin and water soluble carbohydrates on Y_B, while crude lipids and hemicellulose increased biogas formation. For a more diverse substrate (energy crops and animal manure), Triolo et al. (2011) suggested a Y_M prediction model with lignin content as the only one regressor, which reduced Y_M with increasing lignin content. Moreover, Thomsen et al. (2014) confirmed a strong negative influence of lignin on Y_M. However, statistically the content of a chemical compound is related to its effect on Y_M and not necessarily to its properties. The contradictory results published so far

* Corresponding author. Tel.: +49 8161 71 3792; fax: +49 8161 71 4363.
E-mail address: vasilis.dandikas@lfl.bayern.de (V. Dandikas).

http://dx.doi.org/10.1016/j.biortech.2014.10.019
0960-8524/

and the lack of studies addressing the diversity of agricultural plants for biogas production motivated this study. The aim of this study was to develop a model to estimate the Y_B and Y_M that is applicable across different plant species.

2. Methods

2.1. Substrate

To assess the influence of plant species' chemical composition on Y_B, a total of 41 samples of 11 different crops were selected. All samples were gently dried at 40 °C to prevent the ensiling process. Fodder analysis was carried out according to the European regulations (Commission Regulation, 2009) and to the methods of Association of German Agricultural Analytic and Research Institutes (VDLUFA, 1976). Total solids (TS), crude ash (XA), crude lipids (XL), crude fibers (XF), starch (ST), water soluble carbohydrates (WSC), neutral detergent fibers (NDF), acid detergent fibers (ADF), and acid detergent lignin (ADL) were measured; volatile solids (VS), crude protein (XP), nitrogen-free extract (XX), hemicellulose (HC), cellulose (CL), non-fiber carbohydrates (NFC), and organic residue (OR) were calculated.

Table 1
Chemical composition, biogas and methane yield of all 41 samples. The randomly selected 10 samples for model validation are highlighted in gray.

Crop	TS [%FM]	VS [%FM]	XP [%TS]	XL [%TS]	XF [%TS]	ST [%TS]	WSC [%TS]	NDF [%TS]	ADF [%TS]	ADL [%TS]	Y_B [L/kg$_{VS}$]	Y_M [L/kg$_{VS}$]
Barley 1	91.2	86.0	7.7	2.0	22.3	8.4	13.5	50.2	28.2	2.5	663	337
Barley 2	90.6	86.3	7.3	1.8	17.5	11.4	18.1	43.2	22.3	1.8	617	311
Clover 1	92.0	81.9	12.5	3.0	21.5	ND	14.1	47.9	28.5	2.4	619	323
Clover 2	88.9	79.4	14.9	3.4	19.4	ND	13.4	45.1	26.3	1.9	713	370
Clover 3	88.0	80.0	20.1	2.5	13.3	13.4	8.3	32.5	20.6	4.2	531	280
Cup plant 1	91.8	81.0	9.0	4.4	23.8	3.5	2.8	47.0	41.7	6.6	443	231
Cup plant 2	90.3	79.4	11.9	4.0	19.8	4.5	2.2	46.7	36.0	5.5	446	233
Grassland 1	90.2	84.7	11.7	2.5	19.0	ND	23.6	42.5	22.4	1.3	701	360
Grassland 2	91.2	84.2	7.6	3.0	32.8	4.0	5.8	55.9	46.3	8.3	431	225
Grassland 3	91.4	85.5	13.9	3.0	22.4	ND	12.5	52.8	27.8	4.0	653	338
Grassland 4	93.9	88.8	9.5	1.3	23.6	6.2	7.2	57.6	32.5	5.0	555	281
Grassland 5	91.3	85.5	8.3	3.6	29.5	4.6	3.9	54.7	49.2	9.8	437	216
Grassland 6	90.1	84.4	7.9	2.9	30.2	6.6	4.6	59.1	47.3	9.3	339	177
Grassland 7	90.4	84.9	11.8	2.9	21.6	ND	15.7	51.3	27.2	2.0	715	371
Grassland 8	92.3	86.3	8.8	2.1	22.7	6.2	3.6	59.5	39.7	10.8	356	180
Grassland 9	89.7	82.8	7.1	4.5	23.7	5.5	11.7	47.1	35.9	5.8	490	253
Maize 1	90.5	87.9	7.6	3.7	17.8	26.1	6.2	40.0	25.1	2.4	733	369
Maize 2	91.1	87.5	8.5	3.4	13.3	34.0	6.4	35.5	18.7	1.2	799	401
Maize 3	90.1	87.7	7.1	3.2	19.9	27.0	5.6	45.0	26.6	3.3	706	360
Maize 4	90.3	87.4	7.8	3.4	15.4	33.6	4.8	42.1	21.3	1.8	640	327
Millet 1	91.2	86.4	8.5	2.5	26.4	11.6	6.1	57.7	36.6	5.9	514	264
Millet 2	91.8	86.7	8.1	1.4	20.7	18.0	15.0	43.8	27.2	2.4	665	328
Millet 3	92.7	88.1	8.4	2.0	28.2	1.1	10.0	66.7	39.1	5.8	464	238
Millet 4	90.2	84.3	9.9	2.5	19.0	12.8	13.5	47.2	28.7	4.4	667	342
Potato 1	89.2	84.4	8.3	0.3	4.5	66.8	1.5	10.9	7.1	0.1	746	362
Potato 2	88.5	83.3	6.6	0.3	3.8	70.9	1.7	8.3	5.8	0.1	700	352
Rye 1	90.2	86.1	6.5	1.7	22.9	16.2	13.0	45.1	29.3	2.9	629	318
Rye 2	93.1	86.1	10.0	1.8	28.9	1.6	8.2	62.0	35.8	4.0	689	347
Rye 3	92.4	88.6	6.1	1.5	27.2	2.6	15.2	55.4	34.2	3.2	713	352
Rye 4	90.7	84.6	10.0	2.3	27.1	ND	12.7	58.5	33.7	3.7	642	327
Sugar beet 1	94.4	88.3	3.9	0.1	3.8	0.5	72.8	11.0	5.6	0.3	700	350
Sunflower 1	91.5	81.8	9.8	13.3	19.2	3.9	9.0	40.8	30.1	3.7	519	285
Sunflower 2	91.4	82.8	10.8	12.1	18.7	6.5	8.6	42.3	32.1	5.0	621	340
Triticale 1	90.4	82.4	12.6	1.9	26.0	4.4	9.3	52.2	30.9	2.5	680	351
Triticale 2	92.2	87.2	10.3	1.4	29.3	5.3	7.0	62.7	37.3	5.1	617	313
Triticale 3	91.4	85.1	8.1	1.5	31.8	ND	10.3	64.2	37.6	3.5	622	316
Triticale 4	91.3	87.3	6.1	2.4	24.8	9.5	11.9	54.3	33.1	4.1	630	303
Triticale 5	91.1	83.6	11.8	1.7	31.0	2.5	5.2	63.6	37.8	3.5	602	310
Triticale 6	90.5	86.1	6.6	2.1	25.0	6.1	13.5	55.1	32.3	3.1	751	378
Triticale 7	91.5	86.5	10.0	1.7	29.1	5.7	7.4	62.8	38.3	5.4	597	310
Triticale 8	89.7	83.9	8.6	1.6	26.0	6.3	11.0	56.8	35.3	3.4	648	329
MIN	88.0	79.4	3.9	0.1	3.8	0.5	1.5	8.3	5.6	0.1	339	177
MAX	94.4	88.8	20.1	13.3	32.8	70.9	72.8	66.7	49.2	10.8	799	401
SD	1.3	2.5	2.9	2.5	7.1	15.9	11.0	13.6	9.9	2.5	112	55
Median	91.1	85.5	8.5	2.4	22.7	6.3	9.0	50.2	32.1	3.5	630	327
Mean	91.0	85.0	9.3	2.8	22.0	13.2	10.9	48.2	30.5	4.0	610	311
CV	1%	3%	31%	88%	32%	121%	101%	28%	32%	63%	18%	18%

ND: not delectable

2.2. Inoculum for the batch test

The effluent of a pilot biogas plant (working volume of 2.5 m^3) was used as inoculum for all batch experiments. The pilot biogas plant was located in Freising, Germany. The digester was fed with 80% cattle manure and 20% of a dairy cattle feeding mixture (mostly maize and grass silage) at an organic loading rate of 3.0 $kg_{VS}/(m^3 * d)$ and a hydraulic retention time of 19 days at 38 ± 1 °C.

2.3. Batch test

In order to determine Y_B and Y_M, batch experiments were performed based on the German technical guideline VDI 4630 (2006). The batch digester had a working volume of approximately 1.5 L and each sample was tested in triplicate. The ratio of organic dry matter of the sample to organic dry matter of the inoculum was 0.5 ± 0.1. The batch test was conducted at 39 ± 0.5 °C. The volume of biogas was measured via milligas counter (Ritter Apparatebau GmbH, Bochum, Germany). The produced biogas from all three replicates of one sample was stored in a gas bag. A gas analysis automatically took place for aliquots of 1.5 L of produced gas. The gas analysis was performed using an infrared sensor for methane und carbon dioxide measurement and by an electrochemical sensor for oxygen measurement (Awite Bioenergie GmbH, Langenbach, Germany). The Y_B and Y_M potentials were reported as standard liter (dry gas at 273.15 K and 1013.25 mbar) for each kilogram volatile solids (L_N/kg_{VS}).

2.4. Statistical analysis

Descriptive statistics using simple and multiple linear regressions were performed to develop models to predict the potential Y_B. Furthermore, to reduce the interrelated effect of the variables, a principal components analysis (PCA) was conducted. To evaluate and compare the models, the parameters coefficient of determination (R^2), root mean square error (RMSE), normalized RMSE (NRMSE), and coefficient of variation of the RMSE (CVRMSE) were used. The software SAS 9.3 (SAS Institute, USA) and Unscrambler 10.3 (CAMO Software, Norway) were used for the statistical analysis. For the calibration of the developed model, a total of 31 samples were used. In addition, to validate the model 10 samples were randomly selected which were not used during the model calibration (Table 1).

3. Results and discussion

3.1. Biogas yield potential and chemical composition

Based on the batch experiments, the measured Y_B ranged from 339 to 799 L_N/kg_{VS} and the Y_M from 177 to 401 L_N/kg_{VS} (Table 1). The high differentiation of the biogas yields is likely due to the different chemical composition of the feed stocks used. Based on chemical composition of the samples, WSC, ST and XL exhibited the highest variation among all parameters. Additionally, ST and XL values were not normally distributed. These findings suggest that these parameters are not suitable variables for a linear regression analysis.

3.2. Correlation between variables and simple regression analysis

Since a strong positive correlation between Y_B and Y_M was noted for all samples ($r = 0.99$, $p < 0.001$) (Table 2), the statistical analysis focused on understanding the impact of key parameters on Y_B. Y_B was negatively correlated with ADL with a correlation coefficient of −0.90 ($p < 0.001$). In the monocausal regression model ($R^2 = 0.80$, RMSE = 55, NRMSE = 12%, CVRMSE = 9 %), ADL reduced Y_B starting with an intercept value of 775 L/kg_{VS} (Eq. (1)). Y_M was obtained by multiplying the predicted Y_B with a factor of 0.51. Eq. (2) describes the equivalent expression for Y_M:

$$Y_B[L/kg_{VS}] = 775 - 3.93\ ADL\ [g/kg_{VS}] \quad (1)$$

$$Y_M\ [L/kg_{VS}] = 395 - 2.00\ ADL\ [g/kg_{VS}] \quad (2)$$

The model of Triolo et al. (2011) ($Y_M = 460.6–2.58$ ADL, $n = 10$) indicates a stronger impact of lignin and an intercept that was approximately 65 L higher (reflecting higher values for lignin-free fibrous biomass than in this study). This difference may be explained through the lower lignin content (1–6% TS) of the samples used in the Triolo et al. (2011) study.

Statistically, ADL was the most suitable variable for a monocausal model across all energy crops species investigated. However, the accuracy of the model needs improvement to become a useful tool in practice.

Table 2
Correlation matrix of biogas yield, methane yield and the chemical compounds (extended Weender – Van Soest analysis) of all samples.

	XL	XF	XX	ST	WSC	HC	CL	ADL	OR	NFC	YB	YM
XP	0.15	-0.02	-0.39	-0.33	0.18	0.08	-0.02	0.02	0.25	-0.27	-0.11	-0.04
XL		-0.05	-0.36	-0.22	0.00	-0.29	0.13	0.15	0.27	-0.19	-0.19	-0.08
XF			-0.85	-0.83	0.19	0.60	0.95	0.59	0.17	-0.90	-0.43	-0.42
XX				0.91	-0.23	-0.44	-0.86	-0.57	-0.33	0.94	0.48	0.41
ST					-0.50	-0.60	-0.82	-0.50	-0.40	0.94	0.38	0.33
WSC						0.48	0.06	-0.30	0.04	-0.25	0.34	0.36
HC							0.39	0.00	-0.36	-0.66	0.16	0.16
CL								0.71	0.36	-0.88	-0.59	-0.57
ADL		|0| - |0.5| no or low correlation							0.42	-0.60	-0.90	-0.89
OR		|0.5| - |0.8| moderate correlation								-0.20	-0.44	-0.42
NFC		|0.8| - |1| high correlation									0.46	0.41
YB												0.99

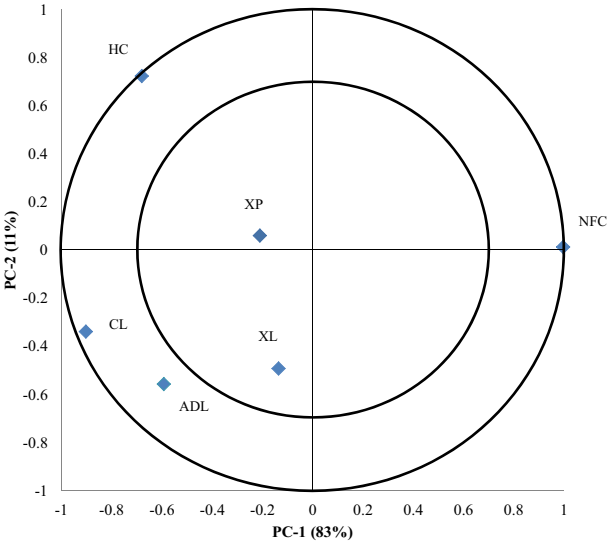

Fig. 1. Correlation loadings of plant chemical compounds (PC1, PC2). The outer circle indicates 100% explained variance and the inner circle indicates 50% explained variance.

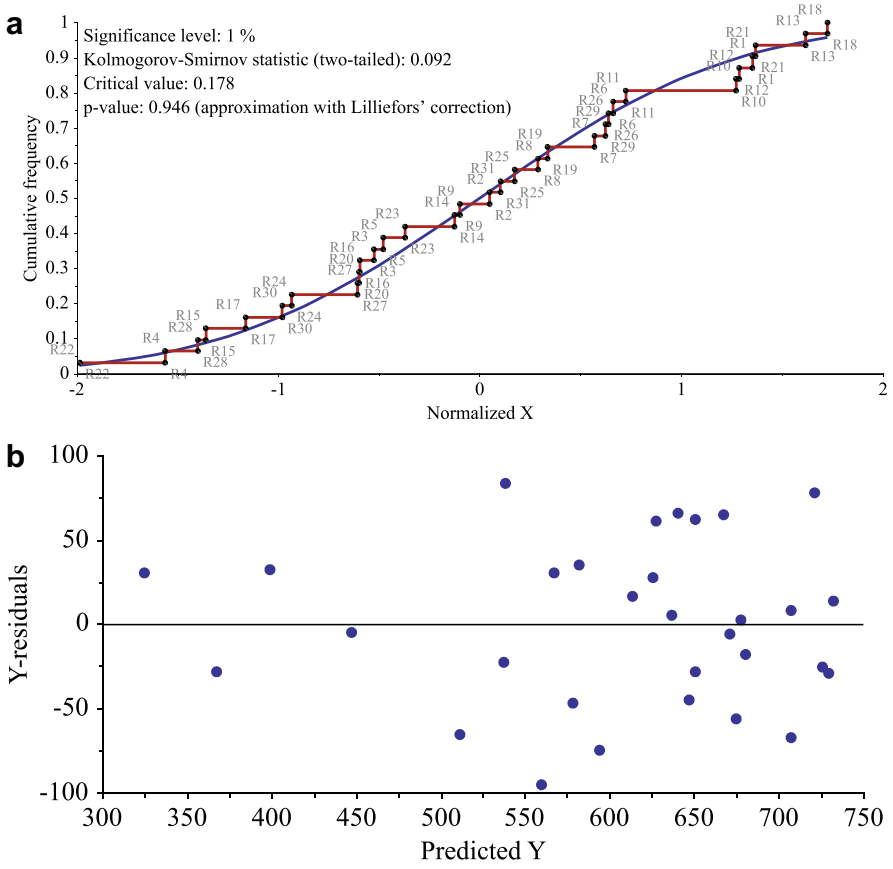

Fig. 2. Kolmogorov–Smirnov test for normal distribution of the residuals of Eq. (3) at significance level of 1% (a) and the residuals of the model plotted versus the predicted values (b).

3.3. Principal component analysis (PCA)

The experimental data were subject to a PCA where the first two principal components (PCs) explained 94% of the variance (Fig. 1). XP and XL are located very close to the center of the plot, i.e. they did not explain the variation of the samples. Therefore, these parameters do not represent suitable variables for a regression model. ADL, HC, CL and NFC are the parameters that mainly explained the variation of the samples and thus can be potentially useful variables for a regression analysis. Based on the correlation matrix (Table 2), only HC was statistically independent from ADL ($r = 0.003$). In contrast, CL and NFC were highly correlated to ADL and hence these are not resilient as a second regression parameter.

3.4. Multiple linear regression analysis

A MLR analysis with ADL and HC as regressors and Y_B as regressand was performed. Both variables were significant for the model ($p < 0.05$) and the statistical performance parameters were slightly improved ($R^2 = 0.83$, RMSE = 47, NRMSE = 10%, CVRMSE = 8%). The model is described by Eq. (3). The equivalent model for Y_M can be derived by multiplying Y_B times 0.51 (Eq. (4)):

$$Y_B\ [L/kg_{VS}] = 727 + 0.25\ HC\ [g/kg_{VS}] - 3.93\ ADL\ [g/kg_{VS}] \quad (3)$$

$$Y_M\ [L/kg_{VS}] = 371 + 0.13\ HC\ [g/kg_{VS}] - 2.00\ ADL\ [g/kg_{VS}] \quad (4)$$

The residuals of the model have been analyzed with the Kolmogorov–Smirnov test for normality and the null hypothesis could not be rejected (Fig. 2a). Moreover, data illustrated in Fig. 2b suggests that the residuals were unsystematically distributed. This residual analysis indicates that the model has a good ability to account for the variability of the data.

These findings agree with the ones reported by Rath et al. (2013), where HC was a prominent variable in their model as well, and even though ST was present at higher concentration in maize it did not differentiate maize cultivars. The key argument for the irrelevance of ST is that it is easily biologically converted. This may explain why even the extremely high concentration of ST in potato and maize did not affect the regression model.

3.5. Validation

Using an independent data set, the suggested model (Eq. (3)) was validated and resulted in a regression line (measured Y_B versus predicted Y_B) of $y = 0.997\ x - 9.717$ with a R^2 of 0.72. The slope was very close to 1 and the intercept was −9.7, both values were close to unity ($y = x$) indicating a close approximation to the data. A Y_B prediction model based on specific plant species could potentially be more accurate. However, due to high variety of substrate used in agricultural biogas plant, the development of crop specific models would be very costly. To assist in adopting the proposed model for field applications, an independent validation with more samples should be conducted.

4. Conclusion

Based on the fodder analysis of different energy crops, it was revealed that the biogas yield is significantly negatively correlated with ADL. Statistically, it was shown that apart from ADL, only HC

seems to be relevant for the difference in biogas yield. For the estimation of biogas and methane yields across energy crops species and for an ADL concentration below 10% of TS, the following equations are proposed:

$$Y_B = 727 + 0.25\ HC - 3.93\ ADL \quad (5)$$

$$Y_M = 371 + 0.13\ HC - 2.00\ ADL \quad (6)$$

Acknowledgement

We greatly acknowledge funding of this study provided by the Bavarian State Ministry of Food, Agriculture and Forestry and the Bavarian State Ministry of Economic Affairs and Media, Energy and Technology, Project code: EW/12/04.

References

Amon, T., Amon, B., Kryvoruchko, V., Zollitsch, W., Mayer, K., Gruber, L., 2007. Biogas production from maize and dairy cattle manure – influence of biomass composition on the methane yield. Agric. Ecosyst. Environ. 118, 173–182.

Angelidaki, I., Alves, M., Bolzonella, D., Borzacconi, L., Campos, J.L., Guwy, A.J., Kalyuzhnyi, S., Jenicek, P., van Lier, J.B., 2009. Defining the biomethane potential (BMP) of solid organic wastes and energy crops: a proposed protocol for batch assays. Water Sci. Technol. Res. 59, 927–934.

Commission Regulation, 2009. Commission Regulation (EC) No. 152/2009 laying down the methods of sampling and analysis for the official control of feed. Off. J. Eur. Union 54, 1–130.

De Moor, S., Velghe, F., Wierinck, I., Michels, E., Ryckaert, B., De Vocht, A., Verbeke, W., Meers, E., 2013. Feasibility of grass co-digestion in an agricultural digester, influence on process parameters and residue composition. Bioresour. Technol. 150, 187–194.

Koch, K., Drewes, J.E., 2014. Alternative approach to estimate the hydrolysis rate constant of particulate material from batch data. Appl. Energy 120, 11–15.

Koch, K., Wichern, M., Luebken, M., Horn, H., 2009. Mono fermentation of grass silage by means of loop reactors. Bioresour. Technol. 100, 5934–5940.

Lebuhn, M., Liu, F., Heuwinkel, H., Gronauer, A., 2008. Biogas production from mono-digestion of maize silage-long-term process stability and requirements. Water Sci. Technol. 58, 1645–1651.

Li, Y., Zhang, R., Liu, G., Chen, C., He, Y., Liu, X., 2013. Comparison of methane production potential, biodegradability, and kinetics of different organic substrates. Bioresour. Technol. 149, 565–569.

Lübken, M., Gehring, T., Wichern, M., 2010. Microbiological fermentation of lignocellulosic biomass: current state and prospects of mathematical modeling. Appl. Microbiol. Biotechnol. 85, 1643–1652.

Mittweg, G., Oechsner, H., Hahn, V., Lemmer, A., Reinhardt-Hanisch, A., 2012. Repeatability of a laboratory batch method to determine the specific biogas and methane yields. Eng. Life Sci. 12, 270–278.

Rath, J., Heuwinkel, H., Herrmann, A., 2013. Specific biogas yield of maize can be predicted by the interaction of four biochemical constituents. BioEnergy Res., 1–14.

Thomsen, S.T., Spliid, H., Østergård, H., 2014. Statistical prediction of biomethane potentials based on the composition of lignocellulosic biomass. Bioresour. Technol. 154, 80–86.

Triolo, J.M., Sommer, S.G., Møller, H.B., Weisbjerg, M.R., Jiang, X.Y., 2011. A new algorithm to characterize biodegradability of biomass during anaerobic digestion: influence of lignin concentration on methane production potential. Bioresour. Technol. 102, 9395–9402.

VDI 4630, 2006. Guideline VDI 4630 – Fermentation of Organic Materials Characterisation of the Substrate, Sampling, Collection of Material Data, Fermentation Tests. Verein Deutscher Ingenieure, Düsseldorf, Germany.

VDLUFA, 1976. Die chemische Untersuchung von Futtermitteln (Engl.: Chemical Analysis of Feedstuffs). VDLUFA-Verlag, Darmstadt, Germany, Band 3.

Paper II

energy&fuels

Article

pubs.acs.org/EF

Correlation between Biogas Yield and Chemical Composition of Grassland Plant Species

V. Dandikas,*[,†] H. Heuwinkel,[‡] F. Lichti,[†] J. E. Drewes,[§] and K. Koch[§]

[†]Institute for Agricultural Engineering and Animal Husbandry, Bavarian State Research Center for Agriculture, Am Staudengarten 3, 85354 Freising, Germany

[‡]Department of Agriculture and Food Economy, Hochschule Weihenstephan-Triesdorf, Am Hofgarten 1, 85354 Freising, Germany

[§]Chair of Urban Water Systems Engineering, Technical University of Munich, Am Coulombwall 8, 85748 Garching, Germany

ABSTRACT: Although grassland has a significant biomass potential, many different factors can affect the quality of grassland feedstock. Changes in the chemical composition of grassland biomass can lead to a high variation in biogas potential, even within the same plant species. Therefore, four grass species and two legume species were grown in field plots and harvested at sequential stages of maturity in the first three growths. The samples were investigated in order to mathematically describe the relation between their chemical composition and biogas yield. A global biogas yield prediction model for energy crops allowed the differentiation between the samples for biogas yield. However, due to their distinct difference in plant chemical composition, the estimation accuracy was rather low. A model considering lignin, hemicellulose, and crude protein as regressors was most suitable for predicting the biogas yield of grassland plant species, with an accuracy of 31 L/kg volatile solids.

1. INTRODUCTION

Energy generation by biogas production can be an environmentally and economically attractive alternative to using fossil fuels.[1] Consequently, the number of agricultural biogas plants has increased all over the world. While maize appears to be the most appealing energy crop, the monocultivation and monodigestion of maize can lead to potentially adverse environmental, biological, and economic effects.[2,3] A more flexible operation involving the use of different feedstock is desirable.

In temperate regions, grassland represents a low-cost biomass resource for biogas production.[4] However, due to many factors, such as plant species, date of harvest, stage of plant maturity, habitats, and microclimate, a high variation of the biogas yield (Y_B) can occur.[5] Moreover, it has been reported that the use of grass as feedstock for biogas plants could lead to technical and biological difficulties;[6] thus, substrate pretreatment is needed.[7] Consequently, grassland biomass may be an effective source of green energy production, but its characterization should be taken into consideration.

The biomethane potential (BMP) is a useful parameter for feedstock characterization. Determining the BMP experimentally is a laborious and costly process. Additionally, due to a lack of standardization, it is difficult to ensure accuracy and reproducibility of data among different laboratories. Although a number of standard protocols are available,[8] the experimental setup used is not always comparable.[9] Instead, a modeling approach based on empirical data could potentially quickly generate reliable information regarding the Y_B potential of different substrates used in practice. Moreover, process modeling can also be a useful tool for optimized process control strategies.[10]

Several authors have proposed linear regression models to predict the Y_B and methane yield (Y_M) of energy crops. In previous studies, the lignin content of energy crops was strongly negatively correlated with Y_B and Y_M.[11−13] However, since the prediction accuracy of monocausal models is not high enough, multiple regressor models are needed.[14,15]

The first objective of this study was to investigate the accuracy of Y_B prediction of energy crops using a global model. The second objective was to develop a specific model that predicts the biogas potential of grassland samples. The study of Dandikas et al.[11] has demonstrated that the differentiation on Y_B is a function of the plant's chemical composition and that the Y_B of energy crops can be predicted by the analytical parameters of the fodder analysis. This study focused on the question whether a precise prediction of Y_B relies on more strictly selection of samples than presented so far. Grassland samples are composed of several plant species, even from contrasting plant families (e.g., grasses and legumes) at different development stages. Therefore, this study investigated and analyzed the chemical composition and the Y_B of different grassland plant species at various harvest dates during the first three growth cycles to investigate the accuracy of Y_B prediction by a more homogeneous sample selection.

2. MATERIAL AND METHODS

2.1. Plant Material and Chemical Analysis. Six plant species from grassland were selected for the experiment. Four grass species (*Lolium perenne*, *Dactylis glomerata*, *Poa pratensis*, *Festuca pratensis*) and two legume species (*Trifolium pratense*, *Trifolium repens*) were grown in field plots (10 m^2). Moreover, four cultivars were tested from the species *Lolium perenne* (Table 1). A single field plot was used for each plant species. The aim of the crop trial was to create a representative data set that would be suitable for statistical analysis. The trial was performed in 2013 in Pulling, Germany (48°36′ N, 11°71′ E). The altitude of the test plot was 470 m above sea level. In order to gain

Received: December 15, 2014
Revised: October 17, 2015
Published: October 27, 2015

DOI: 10.1021/acs.energyfuels.5b01257
Energy Fuels 2015, 29, 7221−7229

Table 1. Investigated Plant Species Harvested at Different Dates during the First Three Growths in 2013

		harvest dates[a]											
		growth cycle 1					growth cycle 2			growth cycle 3			
scientific name	cultivar	1	2	3	4	5	1	2	3	1	2	3	total samples
L. perenne	Arvicola	May 2 (41)	May 6 (49)	May 8 (52)	May 14 (55)	Jun 6 (65)	Jul 16 (30)	–	–	Sep 9 (29)	–	–	7
	Respect	May 14 (43)	May 16 (49)	May 17 (51)	May 21 (55)	Jun 6 (60)	Jul 18 (30)	–	–	Sep 6 (29)	–	–	7
	Sponsor	May 19 (39)	–	Jun 12 (53)	Jun 15 (56)	Jun 18 (60)	–	–	–	–	–	–	4
	Sirius	May 19 (39)	–	Jun 12 (53)	Jun 15 (56)	Jun 18 (60)	–	–	–	–	–	–	4
D. glomerata	Husar	May 14 (41)	May 16 (45)	May 19 (51)	May 21 (56)	Jun 13 (60)	Jul 16 (30)	–	–	Sep 14 (29)	–	–	7
P. pratensis	Lato	May 6 (41)	May 8 (47)	May 14 (52)	May 17 (56)	Jun 6 (61)	Jul 23 (30)	–	–	Sep 4 (30)	–	–	7
F. pratensis	Preval	May 14 (41)	May 16 (45)	May 19 (51)	May 21 (55)	Jun 13 (60)	Jul 11 (30)	–	–	Sep 14 (30)	–	–	7
T. pratense	Titus	May 16 (51)	Jun 17 (55)	Jun 20 (61)	Jun 21 (65)	–	Jul 26 (55)	Aug 2 (61)	–	Aug 29 (59)	Sep 5 (65)	Sep 14 (69)	9
T. repens	Lirepa	May 16 (51)	Jun 12 (55)	Jun 15 (59)	Jun 17 (65)	–	Jul 5 (61)	Jul 16 (65)	Jul 23 (69)	Aug 29 (61)	–	Sep 5 (69)	9

[a]The "–" indicates sample loss. Within the parentheses is written the BBCH code.

information about the chemical composition of the crops during their development, each crop underwent several harvests. The collection of the samples took place at primary, secondary, and tertiary growth cycles with advancing harvest dates in each growth cycle. At defined phenological development stages based on the BBCH scale, five harvest dates were scheduled during the primary growth cycle and three harvest dates were scheduled during the secondary and tertiary growth cycles. However, the unusual weather conditions in 2013 affected the plant growth. Hence, the harvest dates of the plants were reduced (Table 1). In May and June, precipitation was unusually high and the physiological maturity of the plants was delayed. Moreover, the warm and dry weather in July 2013 accelerated the plant maturation, after which the plants did not properly develop, even under rather normal weather conditions during the subsequent months of August and September.

The plant species were selected by the Institute for Crop Science and Plant Breeding of the Bavarian State Research Center for Agriculture (Freising, Germany). The most important criterion for the selection of plant species and harvest dates was the potential and practical relevance of using grassland biomass as feedstock in an agricultural biogas plant.

Immediately after harvesting the crops, they were dried in a forced-air oven at 40 °C. The samples were then ground with a cutting mill and passed through a 10 mm sieve (Retsch, Haan, Germany) before they were stored at room temperature. Forage analysis (Weender and Van Soest analysis) was carried out according to European regulations[16] and in accordance with the methods of the Association of German Agricultural Analytic and Research Institutes.[17] All chemical analyses were performed in duplicate by the Central Department for Quality Assurance and Analytics of Bavarian State Research Center for Agriculture (Freising, Germany). Total solids (TS), crude ash (XA), crude lipids (XL), crude fibers (XF), starch (ST), reducing sugars (RS), neutral detergent fibers (NDF), acid detergent fibers (ADF), and acid detergent lignin (ADL) were measured. Volatile solids (VS), crude protein (XP), nitrogen-free extract (NfE), hemicellulose (HC), cellulose (CL), nonfiber carbohydrates (NFC), and organic residue (OR) were calculated [VS = TS − (XA × TS/100), XP = 6.25 N, NfE = 100 − XA − XP − XL − XF, HC = NDF − ADF, CL = ADF − ADL, NFC = 100 − XA − XP − XL − NDF, OR = 100 − XA − XP − XL − ST − RS − NDF]. The parameters TS and VS are expressed as percent fresh matter (FM) and all others parameters are expressed as percent total solids (% TS).

2.2. Batch Trial and the Experimental Setup. The source of inoculum for a batch test can strongly influence the microbial community and, consequently, the biodegradation of a substrate.[18−20] Therefore, a defined inoculum was used for all batch tests. A pilot digester has been operated for several years in order to obtain a defined biocoenosis. This digester had a working capacity of 2.5 m^3 and it had been operating under steady-state conditions (the coefficient of variation of 5 days of methane productivity was continuously lower than 5%). The digester was fed with 80% cattle manure and 20% dairy cattle feeding (total mixed ratio, mostly maize and grass silage) based on volatile solids at an organic loading rate of 3.0 kg VS/(m^3 d) at 38 ± 1 °C and a hydraulic retention time of 19 days. The pilot digester was located in Pulling, Germany. Chemical and microbiological analyses were carried out regularly to control the actual status of the biocoenosis (data not shown). One week prior to a batch trial, the digester effluent was stored at batch test temperature without feeding in order to reduce its own biogas potential.[21] The degassed digester effluent was used as inoculum for the batch tests. The average values of the chemical parameters from inoculum were recorded at 3.5 ± 0.1% FM of volatile solids, 192 ± 82 mg/kg FM of volatile fatty acids, 8642 ± 868 mg/kg FM of total inorganic carbon, and a pH value of 7.6 ± 0.1.

The batch experiments were performed according to the German standard procedure.[21,22] The batch digester had a total volume of 2 L and a working volume of approximately 1.4 L. The batch digester was filled with 400 mL of distilled water, 1000 g of FM inoculum, and 20 g of FM sample. The ratio of the volatile solids of the sample to the volatile solids of the inoculum was 0.5 ± 0.1. The batch test ran at 39 ± 0.5 °C. Two control samples were used to check the biological activity of the inoculum. Microcrystalline cellulose and a defined sample of dried whole crop maize were used as reference samples. The inoculum alone served as the blank value. Each sample was tested in triplicate (technical replicates), but microcrystalline cellulose was tested in sextuplicate. The TS content in the batch digester ranged between 4% and 5% of fresh matter. The measurement was terminated when the daily biogas production fell below 0.5% of the total volume of biogas that was sampled until that point in time.[22]

The biogas produced was measured with milligas counters (Ritter Apparatebau GmbH, Bochum, Germany) at an accuracy of ±3%. The

DOI: 10.1021/acs.energyfuels.5b01257
Energy Fuels 2015, 29, 7221−7229

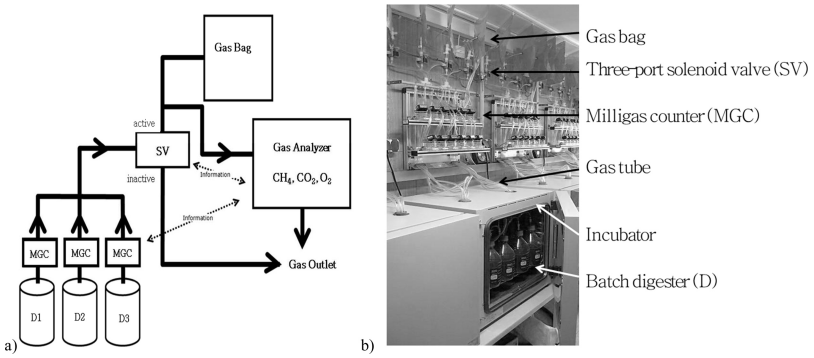

Figure 1. (a) Schematic diagram and (b) photo of the biogas/biomethane potential test system.

measurements were recorded online and the data saved on an hourly basis. After the gas volume determination, the biogas produced from all three replicates of one sample was stored in a gas bag (Figure 1). A gas analysis automatically took place when 1.5 L of gas was collected, i.e., about 0.5 L gas per digester. The gas analyzer individually measured each gas bag. During the experiment, an average of 33 gas analyses was recorded per triplicate of a sample. To perform the gas analysis, an infrared sensor measured methane and carbon dioxide at an accuracy of ±2% and an electrochemical sensor measured oxygen at an accuracy of ±1% (Awite Bioenergie GmbH, Langenbach, Germany). In addition, air pressure and room temperature were recorded on an hourly basis. The Y_B and Y_M were calculated as standard liter (dry gas at 273.15 K and 1013.25 mbar) for each kilogram of volatile solids (L/kg VS).[22] The values of Y_B and Y_M were the average value of three replicates resulting in a coefficient of variation (CV) below 10%.[23]

2.3. Statistical Analysis. Due to the fact that only the organic complex can be utilized during the anaerobic digestion, the content of all chemical compounds were calculated as grams per kilogram of volatile solids (g/kg VS).

The Kolmogorov−Smirnov test was applied at a significance level of 5% to analyze the normal distribution of the samples. Moreover, descriptive statistics and principal component analysis (PCA) were conducted to investigate the correlations among the chemical compounds. The PCA was used to reduce the interrelated effect of the variables, while as much of the variation as possible was retained. Variables that are correlated with each other but are largely independent from other subsets of variables were combined into principal components (PC). Each PC is a linear combination of observed variables and is not correlated (orthogonal) with another PC. The first PC extracts the maximum variability of the observed variables. Sequential PCs are formed from the residual correlations and are orthogonal to all other PCs.[24,25]

Furthermore, multiple linear regression (MLR) analysis was performed in order to develop models predicting the potential Y_B. The parameters coefficient of determination (R^2), root-mean-square error (RMSE), and coefficient of variation of the RMSE (CVRMSE) were used for evaluation and comparison of the models. The software SAS 9.3 (SAS Institute) and Unscrambler 10.3 (CAMO Software) were used for statistical analysis.

3. RESULTS AND DISCUSSION

3.1. Biogas Yield Prediction Using a Global Model. According to the batch trials, the Y_B ranged from 500 to 768 L/kg VS and the Y_M from 263 to 425 L/kg VS with a coefficient of variation of approximately 10% for both variables (Table 2). As expected, a wide range of normally distributed Y_B and Y_M was recorded. The forage analysis of the samples reflected the plant senescence. The different contents of chemical compounds affected the biogas production. Figure 2 displays the chemical composition (XP, NDF, and ADL) and the measured Y_B of the samples obtained from the first growth cycle. It was observed that an increase of slowly biodegradable fractions resulted in a decrease of Y_B. Moreover, ADL was significantly negatively correlated with Y_B (Table 3). However, only around 50% (R^2 = 0.53) of the variance could be explained by ADL. Since the correlation between Y_B and each individual chemical compound was weak, a monocausal regression did not sufficiently explain the variation. Similar observations were reported by Gunaseelan.[26]

Since the Van Soest analysis provides information beyond the Weender analysis and the fractions NDF, ADF, and ADL describe the status of cell wall development, the less complex variables XF and NfE were not used in statistical analysis. Moreover, Y_B was weakly correlated with XF and NfE (Table 3). At this point, it should be considered that although both analytical methods insufficiently describe the cell wall as a biological structure,[27] these analytical parameters can reflect the differentiation of the plant samples' chemical composition and can be thus used for statistical analysis.

For all grass species, no ST was detected, the CV value was extremely high, and the values of this parameter were not normally distributed. Thus, ST could not be used for further statistical analysis. The variations of RS, XP, ADL, and HC were the highest, displaying CV of 43%, 42%, 40%, and 35%, respectively. The highest correlation coefficient (r) between Y_B and these four single variables (RS, XP, ADL, HC) was recorded for ADL and HC, with a r of −0.73 and 0.63, respectively (Table 3). The high variation of these two variables and their good correlation with Y_B potentially explain the variation in Y_B. These findings agree with the results obtained in the study of Dandikas et al.[11] Thus, the across plant species model from Dandikas et al.[11] was utilized for predicting Y_B.

Figure 3 depicts the measured versus predicted data in an *XY* plot. On the basis of slope and offset of the regression line between the two data sets, the model underestimated the Y_B with a bias of −40. Moreover, the RMSE was 75 L/kg VS and the CVRMSE was approximately 12%. Although the estimation error was higher than the estimation error during calibration, the model predicted the Y_B of all samples with a slope of 1.13 (Figure 3). A slope close to 1 demonstrates that the model can predict the range of the measured values and the differentiation between the samples on Y_B. The low accuracy of the model could be explained by the fact that the values for ADL and HC (the two regressors of the model) were not normally distributed. Another reason might have been the fact that for grassland samples the variables HC and ADL alone did not explain sufficiently the Y_B.

DOI: 10.1021/acs.energyfuels.5b01257
Energy Fuels 2015, 29, 7221−7229

Table 2. Chemical Composition, Biogas Yield, and Methane Content of the 61 Samples[a]

sample[b]	TS (% FM)	VS (% FM)	XP (% TS)	XF (% TS)	XL (% TS)	ST (% TS)	RS (% TS)	NDF (% TS)	ADF (% TS)	ADL (% TS)	Y_B (L/kgVS)	CH_4 (%)
G_1	89.0	81.1	17.1	22.0	3.0	ND	17.0	48.5	27.1	3.0	720	53
G_2	90.2	83.5	14.2	21.9	2.5	ND	16.8	49.9	27.0	2.1	768	55
G_3	88.2	80.9	12.4	25.7	2.1	ND	15.8	54.3	31.1	3.8	748	52
G_4	91.3	85.4	9.8	24.5	2.0	ND	18.3	55.1	31.6	2.5	671	55
G_5	91.2	85.9	6.7	27.0	1.6	ND	18.5	60.0	34.3	3.7	646	55
G_6	89.4	82.4	7.2	20.9	2.1	ND	21.9	50.0	28.7	2.8	639	51
G_7	91.8	84.7	15.1	19.7	3.5	ND	15.4	49.8	25.8	5.6	644	53
G_8	89.8	82.2	15.1	23.7	2.7	ND	14.0	53.5	30.5	3.8	732	53
G_9	91.1	84.4	12.1	22.2	2.4	ND	15.5	54.3	30.8	3.2	669	55
G_10	90.3	83.1	13.4	25.1	2.4	ND	15.3	54.6	30.0	2.6	715	52
G_11	90.6	84.4	10.6	24.0	1.9	ND	17.3	53.9	30.0	2.6	675	55
G_12	90.5	85.7	8.3	27.4	1.4	ND	17.2	59.6	33.5	3.2	622	55
G_13	91.5	85.2	7.3	23.8	1.8	ND	15.4	58.0	32.3	3.6	609	52
G_14	94.0	85.8	18.4	20.5	3.6	ND	7.9	53.2	26.2	5.0	599	55
G_15	90.4	82.8	14.6	21.4	2.9	ND	16.7	51.1	27.5	2.4	699	52
G_16	91.2	85.3	7.4	31.1	1.4	ND	15.6	61.9	35.3	3.4	622	55
G_17	91.0	86.4	6.6	26.6	1.4	ND	21.6	57.0	32.3	3.0	631	54
G_18	91.3	86.0	6.7	28.3	1.3	ND	18.6	59.4	34.9	3.8	598	54
G_19	89.4	81.1	14.5	21.8	2.9	ND	16.9	49.5	26.9	2.0	691	52
G_20	91.0	85.7	7.1	27.2	1.6	ND	18.1	57.8	33.0	2.6	630	55
G_21	91.5	87.1	5.9	25.0	1.3	ND	25.7	52.6	30.9	2.7	625	53
G_22	92.1	87.4	5.9	27.9	1.5	ND	20.8	57.4	34.2	3.6	598	54
G_23	90.1	82.1	13.8	28.6	2.7	ND	9.3	62.2	36.3	5.4	633	53
G_24	91.5	84.2	12.2	26.8	2.5	ND	8.5	61.8	35.7	3.1	679	51
G_25	91.2	83.5	11.3	30.7	2.3	ND	10.5	64.5	37.2	3.7	646	53
G_26	91.8	84.6	9.9	27.3	2.7	ND	10.8	62.5	36.0	3.5	655	55
G_27	92.2	86.5	6.6	30.4	1.8	ND	10.8	67.7	40.3	4.4	570	55
G_28	87.4	78.1	9.1	25.2	2.6	ND	10.0	59.4	32.3	4.3	642	53
G_29	93.0	85.3	12.8	24.5	3.0	ND	11.1	57.3	31.4	6.6	697	54
G_30	89.7	82.9	20.3	25.2	2.8	ND	11.6	60.3	29.5	2.7	682	53
G_31	91.5	85.3	18.5	22.4	2.9	ND	9.2	57.7	33.0	4.5	704	56
G_32	90.4	83.8	14.4	30.4	2.4	ND	8.9	68.5	36.7	4.0	642	53
G_33	93.2	87.7	12.6	27.0	2.2	ND	8.7	65.3	35.3	3.6	696	51
G_34	91.8	86.7	9.9	27.6	1.7	ND	10.6	65.3	37.5	4.6	667	54
G_35	90.6	83.6	6.5	21.1	1.6	ND	18.4	59.1	30.3	3.3	717	53
G_36	92.6	86.1	14.6	23.3	2.0	ND	14.4	56.8	29.4	3.1	693	54
G_37	90.1	82.6	14.1	30.3	2.2	ND	8.6	60.5	36.1	3.5	646	51
G_38	92.7	86.3	11.3	27.3	2.0	ND	10.3	59.8	34.5	2.2	687	56
G_39	91.3	84.3	10.8	32.0	1.8	ND	9.9	63.8	37.8	3.3	632	52
G_40	92.1	85.7	9.3	29.6	1.9	ND	10.5	63.9	38.4	3.6	627	55
G_41	91.2	85.9	7.3	30.7	1.7	ND	11.1	66.5	40.0	4.1	604	55
G_42	90.6	82.0	10.7	23.0	2.5	ND	9.9	57.7	31.2	4.3	654	53
G_43	92.5	85.1	14.4	24.2	3.2	ND	7.0	61.2	31.3	5.3	592	55
L_1	88.8	80.5	23.3	19.2	2.6	6.2	9.7	42.7	27.9	6.4	581	54
L_2	90.0	83.6	14.6	25.4	1.9	4.5	10.6	46.0	35.8	6.1	547	56
L_3	89.7	82.3	15.8	25.1	1.9	4.6	7.3	50.6	38.0	6.4	510	56
L_4	89.9	84.5	13.3	27.7	2.7	4.7	8.3	49.6	39.8	6.9	519	56
L_5	88.6	80.8	19.8	19.6	2.5	5.6	7.1	43.5	29.9	6.0	596	53
L_6	89.7	82.3	16.8	18.4	1.8	5.5	6.3	47.7	35.8	6.7	515	53
L_7	91.5	83.3	20.5	19.6	1.9	5.9	7.7	33.6	28.1	5.2	629	55
L_8	91.3	83.4	19.0	21.1	1.8	5.5	8.3	36.4	30.6	5.6	555	54
L_9	91.9	85.1	17.3	22.4	1.5	6.0	9.5	41.2	35.2	7.2	553	54
L_10	88.2	79.6	29.0	15.3	3.0	7.4	7.6	46.3	23.9	8.4	617	52
L_11	89.5	80.2	28.6	23.0	1.9	6.6	2.0	44.6	31.3	7.4	538	56
L_12	88.5	79.8	24.4	17.9	2.3	7.2	7.2	45.3	27.7	6.5	598	52
L_13	89.4	81.7	23.0	19.0	2.5	6.4	7.3	32.4	29.7	7.4	592	56
L_14	90.5	81.5	23.4	23.8	2.0	5.7	4.3	41.9	32.4	8.6	532	53
L_15	89.7	81.2	20.3	19.5	2.1	5.9	5.3	47.4	32.6	7.8	500	53
L_16	89.4	80.9	18.5	20.5	2.0	6.3	5.0	47.7	35.9	8.8	520	53

DOI: 10.1021/acs.energyfuels.5b01257
Energy Fuels 2015, 29, 7221−7229

Table 2. continued

sample[b]	TS (% FM)	VS (% FM)	XP (% TS)	XF (% TS)	XL (% TS)	ST (% TS)	RS (% TS)	NDF (% TS)	ADF (% TS)	ADL (% TS)	Y_B (L/kgVS)	CH_4 (%)
L_17	92.8	85.0	23.7	18.1	2.0	7.3	5.2	42.4	30.2	6.9	542	54
L_18	91.5	83.4	23.6	17.3	1.8	7.0	5.9	34.8	31.4	7.6	560	54
MIN	87.4	78.1	5.9	15.3	1.3		2.0	32.4	23.9	2.0	500	51
MAX	94.0	87.7	29.0	32.0	3.6	7.4	25.7	68.5	40.3	8.8	768	56
CV	1%	3%	42%	16%	25%	158%	43%	16%	12%	40%	10%	3%

[a]For the meaning of the truncations, see the List of Abbreviations. [b]G: Grass sample, L: Legume sample, ND: not detectable.

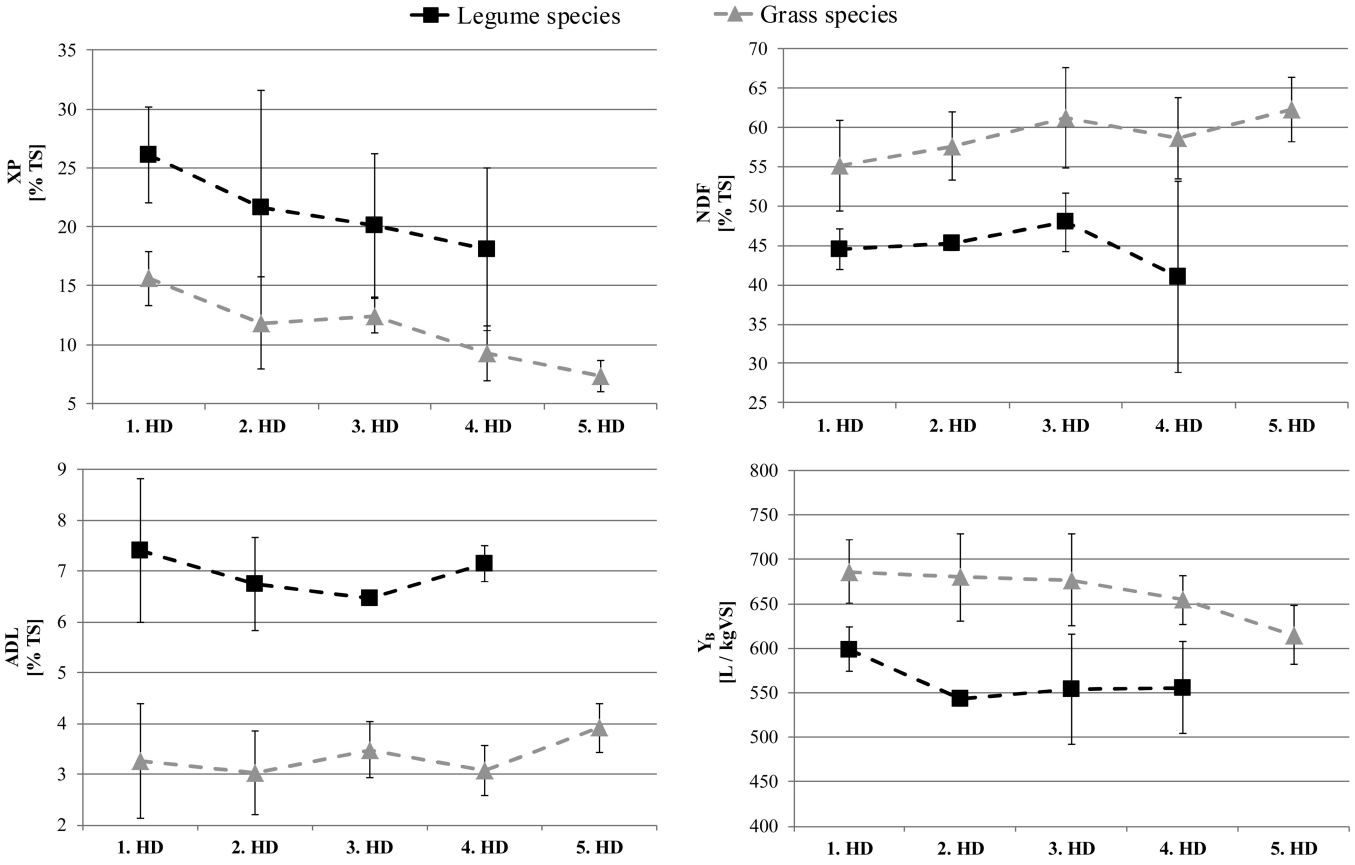

Figure 2. Average values with the standard deviations of crude protein (XP), neutral detergent fiber (NDF), and acid detergent lignin (ADL) content during the plant senescence and the obtained biogas yield (Y_B) during the five harvest dates (HD) of the first growth cycle. Grass samples and legume samples are demonstrated separately.

Table 3. Correlation Matrix of Chemical Compounds Biogas and Methane Yields from the 61 Samples[a]

	XL	XF	NfE	ST	RS	HC	CL	ADL	OR	NFC	Y_B	Y_M
XP	0.41	−0.64	−0.80	0.79	−0.70	−0.57	−0.67	0.72	0.10	−0.11	−0.34	−0.35
XL		−0.30	−0.39	−0.06	−0.26	0.19	−0.44	0.12	−0.30	−0.44	0.27	0.23
XF			0.06	−0.61	0.17	0.52	0.88	−0.47	−0.31	−0.37	0.17	0.19
NfE				−0.51	0.77	0.29	0.21	−0.54	0.15	0.48	0.26	0.27
ST					−0.64	−0.85	−0.49	0.86	0.47	0.29	−0.71	−0.70
RS						0.42	0.09	−0.74	−0.16	0.36	0.53	0.52
HC							0.36	−0.68	−0.80	−0.64	0.63	0.59
CL								−0.41	−0.12	−0.23	0.03	0.06
ADL									0.34	0.05	−0.73	−0.73
OR										0.81	−0.45	−0.38
NFC											−0.23	−0.19
Y_B												0.96

[a]For the meaning of the truncations, see the List of Abbreviations.

DOI: 10.1021/acs.energyfuels.5b01257
Energy Fuels 2015, 29, 7221−7229

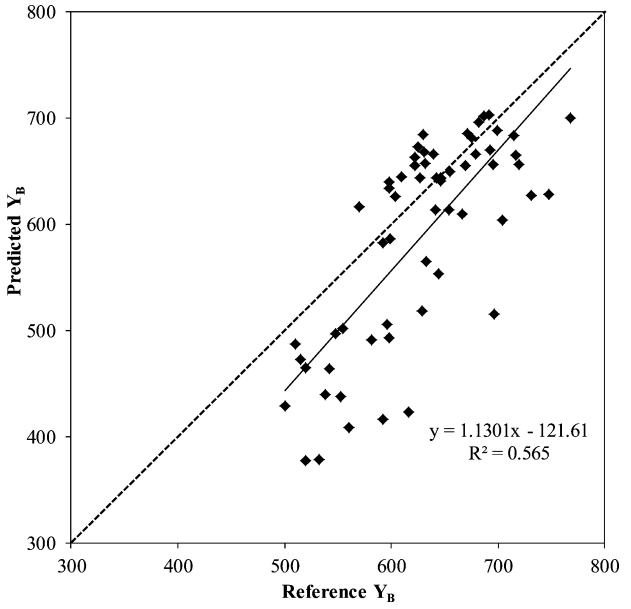

Figure 3. Measured values versus predicted values of biogas yield (Y_B) based on the model of Dandikas et al.

It has been demonstrated earlier for diverse energy crops that only carbohydrates are statistically significant for a Y_B prediction model.[11,13] However, for maize samples, Amon et al.[28] have shown that crude protein and crude lipids are suitable variables for Y_M estimation. Rath et al.[29] reported that crude lipids are statistically significant for the prediction of Y_B of maize cultivars.

Hence, a global model to predict Y_B of typical grassland species is not sufficient, which leads to two hypotheses for the Y_B prediction of grassland samples. First, more than two regressors are needed for a highly accurate model; second, the regressors do not only refer to carbohydrates. Consequently, a plant-specific model may reduce the estimation error.

3.2. Descriptive Statistics. Since a strong positive correlation between Y_B and Y_M was recorded with a r of 0.96 (Table 3) and in order to avoid an additional measurement and its uncertainty, the statistical analysis was focused on the explanation of Y_B. The average methane content of the grassland species was recorded as 54 ± 1.4%.

Although all tested samples are grassland species, they can be classified into two groups based on the plant family. The first group is the grass species and the second group is the legume species. A t-test was conducted to evaluate the difference of the chemical composition between the two plant groups. Figure 4 depicts the average values of the chemical compounds in the two plant groups. A t-test showed that the average XP and ADL contents of grass species were significantly lower than those of the legumes species. Moreover, the average RS, HC, and CL contents of grass species were significantly higher than those of the legumes species. However, no significant difference was found in the average XL content between the two plant groups. For all grass species, no ST was detected. Since during the first growth cycle fewer weather-related sample losses occurred, the t-tests were performed also with the samples only from the first growth cycle (data not shown) in order to check if the number of samples of each plant group affected the t-test results, and similar p-values for each chemical compound were recorded.

As expected, high variance in the chemical composition between the two plant groups was observed. A possible reason for this observation, besides the morphological−physicochemical differences in the two plant groups, could be that the harvest dates and the number of samples were not identical. However, since a mixture of different plant species grow in grassland and are probably at dissimilar morphological development stages for certain harvest dates, the study focused on deriving a global grassland model, which could be a useful tool for practitioners.

3.3. Principal Component Analysis (PCA). For the principal component analysis, the parameters of XP, XL, NFC, HC, CL, and ADL were selected. NFC was selected instead of RS, ST, and OR due to the fact that the variances of these

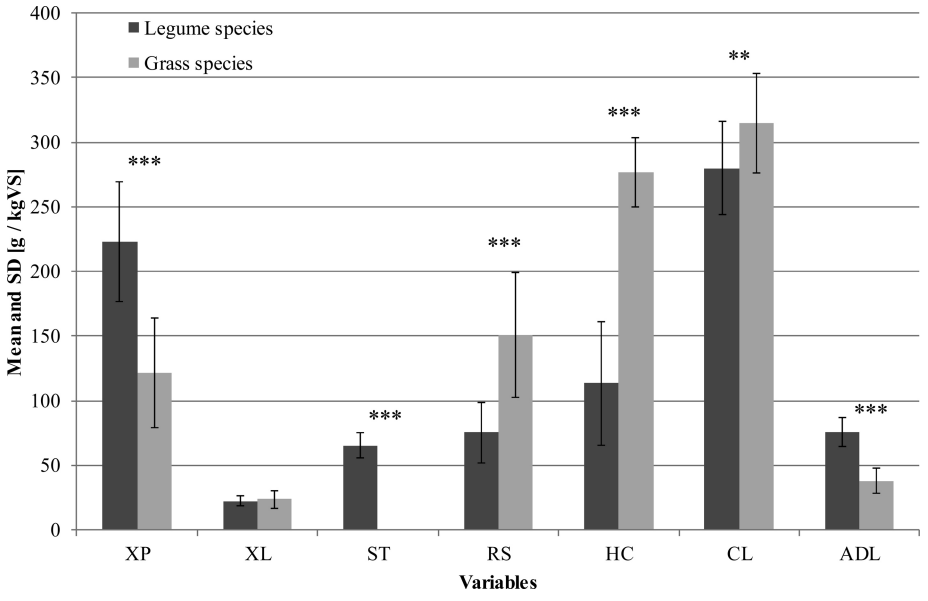

Figure 4. Average values and standard deviations of the chemical compounds of the two plant groups (grass and legume species); p-values were determined by Student's t-test (*p < 0.05, **p < 0.01, ***p < 0.001).

DOI: 10.1021/acs.energyfuels.5b01257
Energy Fuels 2015, 29, 7221−7229

parameters were significantly high and they were not normally distributed. A PCA biplot with the first two principal components is shown in Figure 5. The closer to the outer

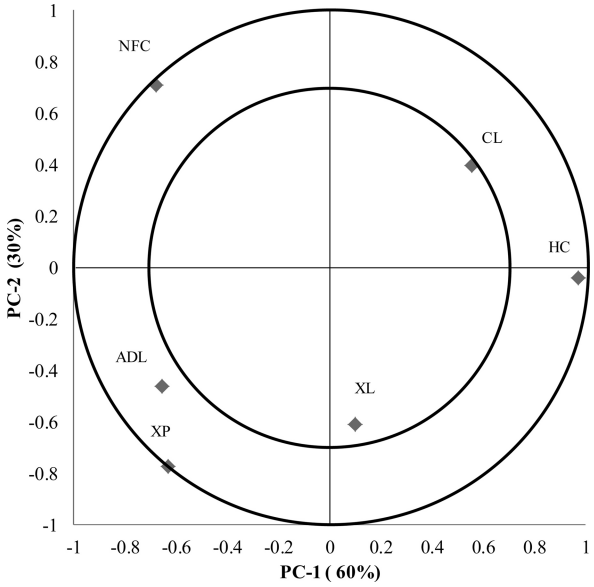

Figure 5. Correlation loadings plot for the first principal component (PC1) versus the second principal component (PC2) of crude protein (XP), crude lipids (XL), nonfiber carbohydrates (NFC), hemicellulose (HC), cellulose (CL), and acid detergent lignin (ADL). The outer circle indicates 100% and the inner circle 50% explained variance.

circle a variable is, the better it can explain the variance. In addition, variables located in the same direction are positively correlated, while variables that are located in opposite directions are negatively correlated. If the angle between two variables is 0° or 180°, they are highly correlated ($r = \pm 1$); if the angle is 90°, they are highly independent.[15,30] The first two PCs explained 90% of the total variation. According to the correlation loadings (Figure 5), XL and CL are not suitable as variables for a regression model. ADL, HC, XP, and NFC are the parameters that mainly explained the variation of the samples and could, therefore, be suitable variables for a regression analysis. Although, XL and CL have been reported as significant variables earlier,[15,29] they were not statistically significant in this study. This was probably due to the fact that XL and CL variations were low in the samples of different growth stages and species.

3.4. Multiple Linear Regression Analysis. An MLR analysis was performed with ADL, HC, XP, and NFC as regressors and Y_B as the regressand. NFC was found not to be significant for the model ($p > 0.05$). Therefore, the MLR analysis was conducted with only three variables (ADL, HC, and XP). During the MLR analysis, three samples were detected as outliers and were excluded from the model calibration. A sample was identified as outlier when the square root of the ratio between the residual calibration variance per sample and the average residual calibration variance for the model was more than 3.0. The analysis of variance (ANOVA) indicated that the model was significant ($p < 0.05$). The R^2 of the model was 0.75; however, the RMSEC was 31 L/kg VS with CVRMSE of 5% (Figure 6). The residuals of the model have been analyzed with the Kolmogorov–Smirnov test for normality at a significance level of 5%, and the null hypothesis could not be rejected (data not shown). Additionally, the

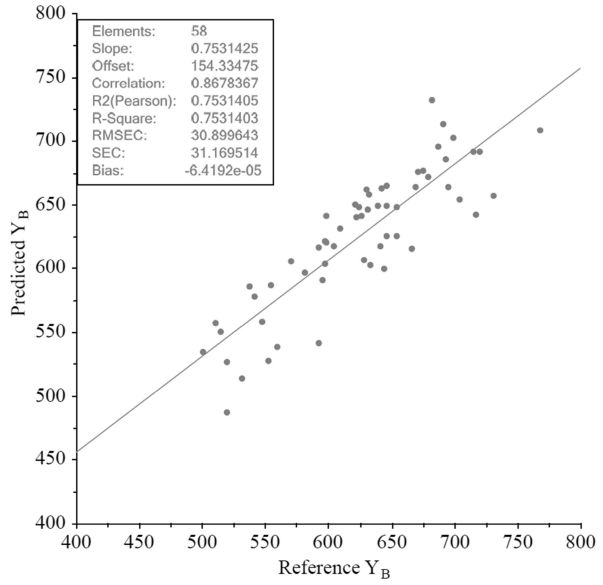

Figure 6. Measured versus predicted values of biogas yield (Y_B) of grassland samples, expressed in L/kg VS. R^2, coefficient of determination; RMSEC, root-mean-square error of calibration; SEC, standard error of calibration.

unsystematic distribution of the residuals indicated that the model was useful for explaining the variability of the data. The model is described by eq 1.

$$Y_B = 670 + 0.44XP + 0.16HC - 3.02ADL \quad (1)$$

where Y_B is expressed in L/kg VS and the chemical compounds are expressed in g/kg VS.

Similar MLR analysis was performed for the estimation of Y_M. The analysis has shown that the R^2 of the model was 0.70, the RMSEC was 18 L/kg VS, and the CVRMSE was 5%. The model is described by eq 2.

$$Y_M = 370 + 0.21XP + 0.05HC - 1.61ADL \quad (2)$$

where Y_M is expressed in L/kg VS and the chemical compounds are expressed in g/kg VS.

The cross-validation has been performed for both models, and similar results were obtained. The results indicated that the XP content of grassland species significantly affected the biogas production of the samples. The correlation matrix showed a weak negative correlation between XP content and Y_B, but according to the MLR, it has a positive influence on Y_B. The monocausal negative correlation between XP and Y_B can be explained by the two plant groups (grass species and legume species). Lower XP values and higher Y_B values were recorded for the grass species, and vice versa for the legume species (Figure 2). The combination of the data led to a mathematical negative correlation. Amon et al.[28] and Gunaseelan[14] have reported a positive effect of XP on Y_B prediction. In contrast, Weissbach[31] has reported that crude protein is not significant for Y_B prediction. Moreover, the protein content decreases during the plant senescence, while the fiber and ADL content increases (Figure 2), attributing to a lower digestibility. Hence, XP depicts plant maturity and it has a positive effect on Y_B. Therefore, eq 1 illustrates that younger plants have higher Y_B.

HC content was the variable with the widest range from 2.7 to 31.8% TS, but in the MLR model it had the lowest contribution, with a regression coefficient of 0.16. The HC was used in MLR to explain the positive effect of the fiber matrix on

DOI: 10.1021/acs.energyfuels.5b01257
Energy Fuels 2015, 29, 7221−7229

biogas production. In agreement with our findings, Amon et al.[28] and Rath et al.[29] reported a positive influence of HC on Y_B prediction.

The ADL content was recorded in a range from 2.0 to 8.8% TS (Table 2), but the effect on Y_B prediction was the highest. Lignin is an amorphous polymer and does not degrade under anaerobic conditions.[32] It is probable that not only lignin alone decreases the Y_B but that the complex lignocellulosic structure also does so, since also a portion of HC and CL is hardly bioavailable. Thus, the regressor ADL in the MLR model reflects the negative influence of the whole fiber matrix. Rath et al.[29] and Triolo et al.[13] have also reported a significant negative influence of ADL on the Y_B prediction.

On the basis of the results of this study, grassland samples with high XP and HC content and low ADL content are preferred feedstock for an agricultural biogas plant.

4. CONCLUSION

The present study revealed that specific models for defined plant groups increase the accuracy of biogas yield prediction. In particular, the developed grassland model reduced the estimation error to 5%. The present model predicts the biogas yield of grassland plant species with an accuracy of 31 L/kg VS using three regressors, namely, acid detergent lignin (ADL), hemicellulose (HC) and crude protein (XP) content. It was shown that ADL reduces biogas yield, while both HC and XP increase it. The regressors of the grassland model reflect the necessity to describe the actual physiological status of the plant and not only its composition. However, plant-specific models lose robustness and increase the risk of misinterpretation if the model is not sufficiently tested to define its range of validity. Consequently, detailed validations with independent data sets should be carried out.

AUTHOR INFORMATION

Corresponding Author

*Tel: +49 8161 71 3792. Fax: +49 8161 71 4363. E-mail: vasilis.dandikas@lfl.bayern.de.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors would like to thank the associate editor and the anonymous reviewers for their invaluable comments and suggestions. This work was funded by the Bavarian State Ministry of Food, Agriculture and Forestry and the Bavarian State Ministry of Economic Affairs and Media, Energy and Technology, Project code BE/14/15.

LIST OF ABBREVIATIONS

ADF = acid detergent fiber
ADL = acid detergent lignin
ANOVA = analysis of variance
BBCH = Biologische Bundesanstalt, Bundessortenamt and Chemische Industrie
BMP = biomethane potential
CH_4 = methane
CL = cellulose
CV = coefficient of variation
CVRMSE = coefficient of variation of root-mean-square error
FM = fresh matter
HC = hemicellulose
HD = harvest date
MLR = multiple linear regression
N = total nitrogen by the Dumas method
NDF = neutral detergent fiber
NFC = nonfiber carbohydrates
NfE = nitrogen-free extract
OR = organic residue
PCA = principal component analysis
r = correlation coefficient
R^2 = coefficient of determination
RMSE = root-mean-square error
RMSEC = root-mean-square error of calibration
SEC = standard error of calibration
ST = starch
TS = total solids
VS = volatile solids
XA = crude ash
XF = crude fibers
XL = crude lipids
XP = crude protein
RS = reducing sugars
Y_B = biogas yield
Y_M = methane yield

REFERENCES

(1) Weiland, P. *Appl. Microbiol. Biotechnol.* **2010**, *85*, 849−860.
(2) Mast, B.; Lemmer, A.; Oechsner, H.; Reinhardt-Hanisch, A.; Claupein, W.; Graeff-Hönninger, S. *Ind. Crops Prod.* **2014**, *58*, 194−203.
(3) Mayer, F.; Gerin, P. A.; Noo, A.; Lemaigre, S.; Stilmant, D.; Schmit, T.; Leclech, N.; Ruelle, L.; Gennen, J.; von Francken-Welz, H.; Foucart, G.; Flammang, J.; Weyland, M.; Delfosse, P. *Bioresour. Technol.* **2014**, *166*, 358−367.
(4) McEniry, J.; O'Kiely, P.; Crosson, P.; Groom, E.; Murphy, J. D. *Biofuels, Bioprod. Biorefin.* **2011**, *5*, 670−682.
(5) McEniry, J.; O'Kiely, P. *Bioresour. Technol.* **2013**, *127*, 143−150.
(6) Thamsiriroj, T.; Murphy, J. D. *Energy Fuels* **2010**, *24*, 4459−4469.
(7) Tsapekos, P.; Kougias, P. G.; Angelidaki, I. *Bioresour. Technol.* **2015**, *182*, 329−335.
(8) Angelidaki, I.; Alves, M.; Bolzonella, D.; Borzacconi, L.; Campos, J. L.; Guwy, A. J.; Kalyuzhnyi, S.; Jenicek, P.; van Lier, J. B. *Water Sci. Technol.* **2009**, *59*, 927−934.
(9) Raposo, F.; Fernández-Cegrí, V.; De la Rubia, M. a.; Borja, R.; Béline, F.; Cavinato, C.; Demirer, G.; Fernández, B.; Fernández-Polanco, M.; Frigon, J. c.; Ganesh, R.; Kaparaju, P.; Koubova, J.; Méndez, R.; Menin, G.; Peene, A.; Scherer, P.; Torrijos, M.; Uellendahl, H.; Wierinck, I.; de Wilde, V. *J. Chem. Technol. Biotechnol.* **2011**, *86*, 1088−1098.
(10) Blesgen, A.; Hass, V. C. *Energy Fuels* **2010**, *24*, 4721−4727.
(11) Dandikas, V.; Heuwinkel, H.; Lichti, F.; Drewes, J. E.; Koch, K. *Bioresour. Technol.* **2014**, *174*, 316−320.
(12) Thomsen, S. T.; Spliid, H.; Østergård, H. *Bioresour. Technol.* **2014**, *154*, 80−86.
(13) Triolo, J. M.; Sommer, S. G.; Møller, H. B.; Weisbjerg, M. R.; Jiang, X. Y. *Bioresour. Technol.* **2011**, *102*, 9395−9402.
(14) Gunaseelan, V. N. *Bioresour. Technol.* **2009**, *100*, 3426−3429.
(15) Xu, F.; Wang, Z.-W.; Li, Y. *Bioresour. Technol.* **2014**, *173*, 168−176.
(16) Commission Regulation. *Off. J. Eur. Union* **2009**, *L54*, 1−130.
(17) Naumann, K.; Bassler, R. Chemical analysis of feedstuffs. In *VDLUFA-Methodenbuch*; VDLUFA-Verlang: Darmstadt, Germany, 1976; Vol. *III*-Futtermittel.
(18) Gu, Y.; Chen, X.; Liu, Z.; Zhou, X.; Zhang, Y. *Bioresour. Technol.* **2014**, *158*, 149−155.

DOI: 10.1021/acs.energyfuels.5b01257
Energy Fuels 2015, 29, 7221−7229

(19) Lopes, W. S.; Leite, V. D.; Prasad, S. *Bioresour. Technol.* **2004**, *94*, 261−266.
(20) Xu, F.; Shi, J.; Lv, W.; Yu, Z.; Li, Y. *Waste Manage.* **2013**, *33*, 26−32.
(21) VDI. *Guideline VDI 4630—Fermentation of Organic Materials. Characterisation of the Substrate, Sampling, Collection of Material Data, Fermentation Tests*; Beuth-Verlang: Düsseldorf, Germany, 2006.
(22) Wulf, S.; Tillman, K.; Bischoff, M. Measurement of Biogas and Methane Yields in Fermentation Tests. In *VDLUFA-Methodenbuch*; VDLUFA-Verlang: Darmstadt, Germany, 2013; Vol. *VII*-Umweltanalytik.
(23) Heuwinkel, H.; Aschmann, A.; Gerlach, R.; Gronauer, A. Die Genauigkeit der Messung des Gasertragspotentials von Substraten mit der Batchmethode (English: The accuracy of the measurement of the gas yield potential of substrates using the batch method). *Bayer. Landesanst. für Landwirtsch. Int. Wiss. Biogas Sci. 2009*; Bayerischen Landesanstalt für Landwirtschaft: Freising, Germany, 2009; Vol. *1*, 95−103.http://www.lfl.bayern.de/mam/cms07/ilt/dateien/biogas_science_bd_1.pdf.
(24) Jolliffe, I. T. *Principal Component Analysis*; Springer: New York, 2002.
(25) Tabachnick, B. G.; Fidell, L. S. *Using Multivariate Statistics*; Pearson Education: Boston, 2012.
(26) Gunaseelan, V. N. *Bioresour. Technol.* **2007**, *98*, 1270−1277.
(27) Jung, H.-J. G. *J. Nutr.* **1997**, *127*, 810S−813S.
(28) Amon, T.; Amon, B.; Kryvoruchko, V.; Zollitsch, W.; Mayer, K.; Gruber, L. *Agric., Ecosyst. Environ.* **2007**, *118*, 173−182.
(29) Rath, J.; Heuwinkel, H.; Herrmann, A. *BioEnergy Res.* **2013**, *6*, 939−952.
(30) Appels, L.; Lauwers, J.; Gins, G.; Degrève, J.; Van Impe, J.; Dewil, R. *Environ. Sci. Technol.* **2011**, *45*, 4173−4178.
(31) Weissbach, F. *Pflanzenbauwissenschaften* **2009**, *13*, 72−85.
(32) Koch, K.; Lübken, M.; Gehring, T.; Wichern, M.; Horn, H. *Bioresour. Technol.* **2010**, *101*, 8158−8165.

NOTE ADDED AFTER ASAP PUBLICATION

This article published November 10, 2015 with an error in the calculations of organic residue in the third paragraph of the Material and Methods section. The corrected version published November 19, 2015.

DOI: 10.1021/acs.energyfuels.5b01257
Energy Fuels 2015, 29, 7221−7229

Paper III

Renewable Energy 118 (2018) 34–42

Contents lists available at ScienceDirect

Renewable Energy

journal homepage: www.elsevier.com/locate/renene

Correlation between hydrolysis rate constant and chemical composition of energy crops

Vasilis Dandikas [a], Hauke Heuwinkel [b], Fabian Lichti [a], Thomas Eckl [c], Jörg E. Drewes [d], Konrad Koch [d, *]

[a] Institute for Agricultural Engineering and Animal Husbandry, Bavarian State Research Center for Agriculture, Am Staudengarten 3, 85354 Freising, Germany
[b] Department of Agriculture and Food Economy, Hochschule Weihenstephan-Triesdorf, Am Staudengarten 1, 85354 Freising, Germany
[c] Department of Experimental Farm, Bavarian State Research Center for Agriculture, Lange Point 12, 85354 Freising, Germany
[d] Chair of Urban Water Systems Engineering, Technical University of Munich, Am Coulombwall 3, 85748 Garching, Germany

ARTICLE INFO

Article history:
Received 23 April 2017
Received in revised form
14 August 2017
Accepted 28 October 2017
Available online 29 October 2017

Keywords:
Biogas potential
Hydrolysis rate constant
Kinetic model
Lignocellulosic biomass
Linear regression

ABSTRACT

Besides biogas yield, the kinetic of biogas production in a biomethane potential (BMP) test also provides important information for feedstock characterization. In this study, fodder analysis and BMP tests with high temporal resolution were performed in order to identify statistical correlations between the hydrolysis rate constant (k_h) and the chemical composition of various energy crops. Different species and cultivars of energy crops were analyzed in order to develop a broadly applicable regression model for the prediction of k_h. Two independent datasets (222 samples in total) were used, one for the calibration of the model and one for its validation. The results indicated that the analytical parameters non-fiber carbohydrates and crude protein were statistically suitable for a multiple linear regression model for the prediction of k_h. Furthermore, a first-order kinetic model and the proposed regression models can be utilized for the prediction of the biogas production in a BMP test. The proposed approach offers a fast and reliable prediction of the biogas production rate and allows a feedstock assessment according to their biogas potential.

1. Introduction

Biogas utilization is a form of sustainable energy production and allows for less dependence on fossil fuels and for greenhouse gas emission reduction [1,2]. Moreover, biogas as part of a cogeneration energy system for electricity and heat production based on renewable resources has high potential, and those cogeneration systems could meet the current energy demand [3,4]. However, the feedstock characteristics and regional production conditions should also be considered [5]. It has been reported that even lignocellulosic-rich biomass has a high biogas potential and if the harvesting technology is appropriate, it could be utilized as a co-substrate to improve methane production of an agricultural biogas plant [6].

Compared to other renewable energy resources, such as solar and wind power, biogas is advantageous because it can be stored and utilized independent of time and place at times of higher energy demand [7,8]. Different concepts for demand-driven flexible operation of a biogas plant have been suggested [9,10] based on either the storage of biogas or on-demand biogas production by means of the feedstocks. However, the demand-driven flexible operation of biogas plants creates new challenges in biogas technology, and in order to define an optimal flexible operational concept, many factors should be taken into account; one of them is feedstock characteristics [11]. Thus, accurate feedstock assessment is required for an efficient biogas plant operation, which is the main goal on anaerobic digestion technology [12]. Biogas potential and the biogas production rate (kinetics) are crucial parameters for substrate characterization.

When complex substrates are applied, such as biowaste or energy crops, it can be assumed that hydrolysis is the rate limiting step during anaerobic degradation [13,14]. Hence, when no microbial inhibition occurs during the anaerobic digestion process, it can be assumed that the rate of the overall process is mainly driven

* Corresponding author.
E-mail address: k.koch@tum.de (K. Koch).

https://doi.org/10.1016/j.renene.2017.10.100
0960-1481/

Acronym

ADF	Acid detergent fiber
ADL	Acid detergent lignin
ANOVA	Analysis of variance
BMP	Biomethane potential
CL	Cellulose
CV	Coefficient of variation
CVRMSE	Coefficient of variation of root mean square error
FM	Fresh matter
HC	Hemicellulose
k_h	Hydrolysis rate constant
MLR	Multiple linear regression
NDF	Neutral detergent fiber
NFC	Non-fiber carbohydrates
NfE	Nitrogen free extract
OR	Organic residue
PCA	Principal component analysis
R^2	Coefficient of determination
RMSE	Root mean square error
RS	Reducing Sugar
ST	Starch
TS	Total solids
VS	Volatile solids
XA	Crude ash
XF	Crude fibers
XL	Crude lipids
XP	Crude protein
Y_B	Biogas yield

by hydrolysis, which can be described by the hydrolysis rate constant k_h [15].

The experimental determination of biogas and methane potential is commonly carried out in biomethane potential (BMP) tests. Although this method is well described in different guidelines [16–18], round robin tests have highlighted a high deviation of biogas yields from the same feedstocks by different laboratories [19]. Therefore, a recently published guideline aims for a stricter standardization in order to increase both inter- and intra-laboratory reproducibility [20]. However, a BMP test is a high-effort and time-consuming process [21], and a precise prediction of biogas yield by the chemical composition of the substrate is desirable [22–25]. The simulation of the anaerobic digestion process as a white-box is extremely demanding [26] due to the high variability of microbes, feedstock specific microbial communities involved, and the different optimal growth conditions [27–29]. In contrast, considering the anaerobic process as a black-box is a simpler approach and allows the simulation of the biological process with less effort [30]. The most commonly applied approach is a first-order model to describe the biogas production rate in a BMP test [16,31].

Although modeling of biogas production has been well studied already [32–34], knowledge on the correlation between the chemical compounds and the hydrolysis rate constant is limited, and studies on the prediction of hydrolysis rate constant of a BMP test are lacking in the peer-reviewed literature so far. Since fodder analyses are well-established standard methods, and extensive databases of feed chemical composition already exist, a model allowing the simultaneous prediction of both biogas yield and biogas production rate based only on the chemical composition to assess various feedstock based on their biogas potential, is desirable for both researchers and biogas plant operators.

Therefore, the objectives of this study were firstly, to analyze the correlation between the hydrolysis rate constant (k_h) and the chemical composition of the feedstock based on a high variety of plant species and cultivars, and secondly, to develop and evaluate a regression model in order to predict k_h based on feedstock chemical composition. BMP tests and fodder analyses were performed, and the data were statistical analyzed to define and quantify the effect of the various plants' chemical compounds on biogas production. In addition, the k_h prediction model was combined with a previously published biogas yield prediction model [22] in order to predict both biogas yield and biogas production rate based on a first-order kinetic model. Finally, the modeling approach has been validated with an independent dataset.

2. Material and methods

2.1. Substrates

Field experiments were carried out in order to collect different plant species and cultivars of energy crops under well-defined conditions. The field experiments were performed in 2011, 2013, 2014 and 2015 in order to include also different weather conditions. The plants were grown at various locations in Bavaria, Germany. In total, twenty-one different plant species were cultivated and systematically harvested at various developmental stages or growth cycles (grassland species). The collected samples were dried in a forced air oven at 40 °C directly after harvest, then ground in a cutting mill (Retsch, SM 200, Germany) to pass a 10 mm sieve and were stored at room temperature. The plant selection and the field experiments were carried out by the Institute for Crop Science and Plant Breeding of the Bavarian State Research Center for Agriculture, Freising, Germany and by the Technology and Support Center, Straubing, Germany.

Two independent datasets were generated based on the harvest year: one dataset for the calibration of the model with 131 samples of 13 plant species (the samples were harvested in 2011, 2013, 2014), and one dataset for the validation of the model with 91 samples of 10 plant species (the samples were harvested in 2015). All 222 samples were collected and analyzed to assess the influence of plants' chemical composition on biogas potential and on biogas production rate (hydrolysis rate), respectively (see Table SM-1 and SM-2 in the supplementary material (SM)).

2.2. Inoculum and BMP test

Several studies have confirmed that the biological, physical, and chemical characteristics of inoculum can affect the results of a BMP test and moreover the hydrolysis rate [35–38]. It has been reported that the origin of an inoculum for BMP test has a significant effect on the biogas production rate [38]. In order to obtain a well-defined inoculum for the BMP tests, a digester of 2.5 m^3 had been operating at an organic loading rate of 3 $kg_{VS}/(m^3 * d)$ and at mesophilic temperature. More details about the source of inoculum can be found in [39].

In total, eight BMP trials have been carried out to analyze all samples of the calibration's dataset. Table 1 provides the characteristics of the inoculum prior to the experiments. As indicated by the values, the inoculum quality was very similar during the eight independent experiments allowing a comparison of the hydrolysis

Table 1
Inoculum characterization of the eight trails of the calibration's dataset.

Parameter	Unit	T1	T2	T3	T4	T5	T6	T7	T8
TS	[% FM]	4.65	4.98	4.83	4.80	4.53	4.43	4.11	4.44
VS	[% FM]	3.32	3.58	3.57	3.53	3.20	3.08	2.82	3.04
TIC[a]	[g/kg]	8.51	7.46	9.30	9.30	8.96	9.17	9.54	10.0
VFA[b]	[g/kg]	0.39	0.11	0.21	0.15	0.11	0.45	0.25	0.31
NH_4-N	[g/kg]	1.60	1.78	1.93	1.68	1.34	1.51	1.61	1.58
pH	[−]	7.5	7.5	7.7	7.6	7.8	7.6	7.9	7.8
Biogas rate[c]	[mL/(L*d)]	213	313	304	321	234	244	244	221

[a] Total inorganic carbon.
[b] Volatile fatty acids in g of acetic acid equivalent.
[c] Daily biogas production rate of the inoculum, 24 h before the BMP test.

rate of the samples not only within a trial but also among all trials. Moreover, the biogas potential of the inocula (residual gas potential) was similar during the BMP tests with an average value of 104 ± 12 L/kg_{VS} (data not shown).

The BMP tests were performed according to the guidelines of VDI 4630 and VDLUFA [17,18]. Three technical replicates for each sample were used to determine the biogas potential. Moreover, microcrystalline cellulose and dried whole-crop maize as an internal standard were used as positive controls in each trial to check the activity of the inoculum. The biogas quality was analyzed every 1.5 L gas produced; methane and carbon dioxide concentrations were detected with infrared sensors, and oxygen concentration was detected with an electrochemical sensor (Awite Bioenergie GmbH, Langenbach, Germany). The experiments were conducted at mesophilic temperature (38 ± 1 °C). More details about the experimental setup can be found in [39]. The biogas yields are reported as liter of dry gas at standard temperature (273.15 K) and pressure (1013.25 mbar) per kilogram volatile solids added (L/kg_{VS}).

2.3. Chemical analysis

Total solids (TS) and volatile solids (VS) were determined gravimetrically according to German Standard Methods for the Examination of Water, Wastewater and Sludge [40]. Crude fiber (XF) content was determined by the Weender analysis; moreover, the important fiber fractions (NDF, ADF, ADL) were determined by the Van Soest method [41]. Hemicellulose (HC) and cellulose (CL) contents were calculated (HC = NDF - ADF, CL = ADF - ADL). Nitrogen content was determined by the Dumas method and crude protein (XP) was calculated based on the nitrogen content multiplied by 6.25. Crude lipid (XL) content was determined by the extraction method. Starch (ST) content was determined polarimetrically and reducing sugar (RS) content volumetrically. Non-fiber carbohydrates (NFC), and organic residue (OR) were calculated (NFC = 100 - XA - XP - XL - NDF, OR = 100 - XA - XP - XL - ST - RS - NDF). All analytical methods were executed as described by Naumann and Bassler [42] and they were carried out by the Department of Quality Assurance and Analytics of the Bavarian State Research Center for Agriculture, Freising, Germany.

2.4. Statistical analysis

The software SAS 9.3 (SAS Institute, USA) and Unscrambler 10.3 (CAMO Software, Norway) were used for the statistical analysis. Correlation analysis was performed based on Pearson's method. Furthermore, principal components analysis (PCA) was conducted to determine the main explanatory variables for the variation of the dataset. PCA can expose and visualize correlations within the dataset. Similarities and differences of the samples can be revealed plotting the measured variables in a plot of mathematically defined variables called principal components (PC). The first PC extracts the maximum variability of the observed variables. Sequential PCs are formed from the residual correlations and are orthogonal to all other PCs [43]. Consequently, multiple regression analysis was performed to develop prediction models. Moreover, the Mallows' C_p model selection method has been applied in order to find out the best fit model [33]. A low C_p value indicates a good fit of the model. Hence, the lower the C_p value is, the higher the precision of the model will be. The parameters coefficient of determination (R^2), root mean square error (RMSE), and coefficient of variation of the RMSE (CVRMSE) were applied to evaluate and compare the models.

2.5. Estimation of biogas production rate and yield

The cumulative biogas production of energy crops in BMP tests is often described by a simple first-order approach [16,31,44].

$$Y_{B(t)} = Y_B * \left(1 - e^{-k_h * t}\right) \quad (1)$$

Where $Y_{B(t)}$ is the cumulative biogas production at time t in L/kg_{VS}, Y_B is the biogas yield in L/kg_{VS}, and k_h is the first-order hydrolysis rate constant in $h^{-1}$and t is the time in h.

Based on Eq. (1), the hydrolysis rate constant (k_h) can be determined from the slope of the curve after linearization when plotting the logarithm versus time [16]. This approach of hydrolysis rate constant determination is called k_{h_ln} in this study.

To estimate the biogas yield, the model of Dandikas et al. [22] has been chosen. The prediction is based on the content of both hemicellulose (HC) and lignin (ADL):

$$Y_B = 727 + 250\,HC - 3930\,ADL \quad (2)$$

where Y_B is expressed as L/kg_{VS}, and the chemical compounds are expressed as kg/kg_{VS}.

3. Results and discussion

3.1. Calibration of the k_h prediction model

3.1.1. Correlation of chemical compounds, biogas yield and hydrolysis rate constant

Pearson's correlation coefficient (r) was used in order to express the mono correlation between the parameters (Table 2). Among the parameters, a high correlation not only within the carbohydrates but also among protein, lipids and carbohydrates was observed. This can be explained by the plant growth process and the proportional distribution of the chemical compounds, as they are expressed as percentage of TS. A strong negative correlation between biogas yield and ADL was observed also for this dataset, which is in agreement with previous studies [22,24,45,46].

According to Table 2, k_{h_ln} was not or weakly correlated with all chemical compounds, which is in line with observations reported previously [45]. A possible explanation of this observation is that, since linearization was done based on the whole duration of the experiment, the regression line met the linearized curve approximately in the middle of the time axis. However, approximately 80% or more of the total biogas production of energy crops is usually recorded at the first half of the experiment. Thus, the variance of Y_B in the middle of the experiment is similar to the variance of Y_B at the end of the experiment in this dataset of energy crops.

If k_h would be determined at the time period between the start of the experiment and the half-maximum biogas production, then the variation of the dataset may be better explained. Therefore, the hydrolysis rate constant at 50% of Y_B has been defined as alternative to k_{h_ln} as follows:

Table 2
Cross-correlation matrix of chemical compounds, biogas yield and hydrolysis rate constant of the 131 samples of the calibration's dataset.

	XL	XF	NfE	ST	RS	HC	CL	ADL	OR	NFC	Y_B	$k_{h\_0.5}$	k_{h_ln}
XP	0.49	-0.35	-0.66	-0.06	-0.16	-0.47	-0.36	0.45	0.10	-0.12	-0.17	0.52	0.37
XL		-0.30	-0.33	0.00	-0.04	-0.29	-0.26	0.10	0.04	-0.01	0.04	0.30	0.40
XF			-0.46	-0.62	-0.25	0.65	0.95	0.21	-0.08	-0.79	-0.26	-0.68	-0.22
NfE				0.55	0.34	-0.06	-0.42	-0.58	-0.03	0.74	0.35	0.05	-0.20
ST					-0.26	-0.43	-0.63	-0.21	-0.11	0.70	0.22	0.49	0.02
RS						-0.05	-0.29	-0.37	-0.24	0.34	0.38	0.08	0.09
HC							0.56	-0.34	-0.49	-0.68	0.10	-0.72	-0.15
CL								0.26	0.06	-0.76	-0.34	-0.67	-0.23
ADL									0.29	-0.31	-0.70	0.05	-0.02
OR		|0| - |0.5| no or weak correlation								0.22	-0.34	0.10	-0.15
NFC		|0.5| - |0.8| moderate correlation									0.31	0.56	0.01
Y_B		|0.8| - |1| strong correlation										0.24	0.24
$k_{h\_0.5}$													0.45

$$k_{h\_0.5} = \frac{\ln(2)}{t_{0.5}} \quad (3)$$

where $k_{h\_0.5}$ is the hydrolysis rate constant at 50% of Y_B in h^{-1}, and $t_{0.5}$ is the time when half of Y_B has been reached in h.

This approach follows the logic of Koch and Drewes [15], who suggested a Monod-type instead of the first-order model for predicting the BMP of complex substrates.

From Table 2 it can be seen that $k_{h\_0.5}$ was moderately correlated with XP, XF, HC, CL and NFC. This result already confirmed the hypothesis that k_h determined at the time period between the start of the experiment and half-maximum of biogas production ($k_{h\_0.5}$) better describes the kinetics of the biogas production. Therefore, only $k_{h\_0.5}$ was used as dependent variable for the subsequent regression analysis. However, Y_B was only weakly correlated with both hydrolysis constants. This leads to the conclusion that the amount of biogas produced is apparently not directly related to the rate of the anaerobic process. ADL was also not correlated with the two hydrolysis rate constants k_{h_ln} and $k_{h\_0.5}$. There might be two reasons for that: Firstly, lignin is indigestible under anaerobic conditions [47] and its solubilization during the hydrolysis has no impact on the subsequent biogas production. Secondly, the amount of ADL in the dataset was relatively low with an average of 4% and the enzymatic disintegration of the lignocellulosic matrix could be quickly performed.

The values of $k_{h\_0.5}$ were plotted in a histogram (Fig. 1) and were apparently normally distributed around the mean value. This implies that the hypothesis of normal distribution cannot be rejected, which was also confirmed by a Kolmogorov-Smirnov test at a significance level of 0.05. This result indicates that the $k_{h\_0.5}$ is suitable for further statistical analysis and could be used as regressand.

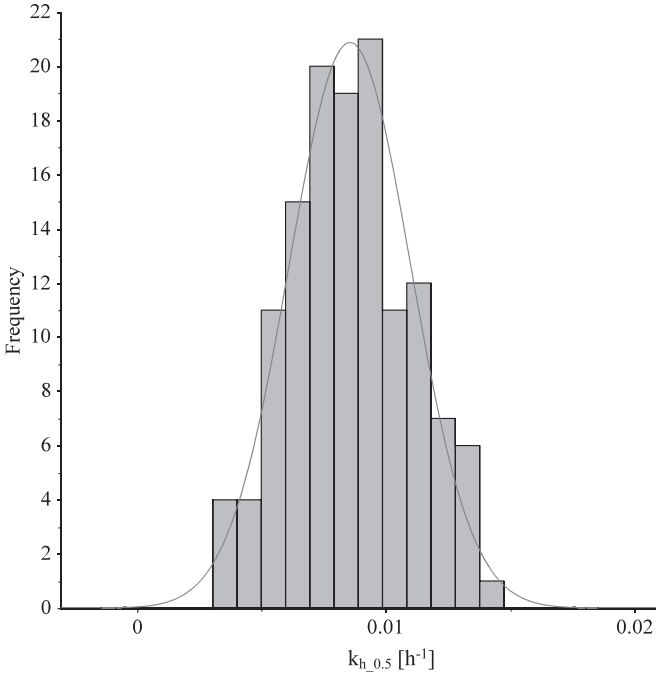

Fig. 1. Histogram for $k_{h\_0.5}$ values of calibration's dataset.

3.1.2. Principal component analysis

In order to examine the correlations among those parameters, a principal component analysis (PCA) with principal component regression (PCR) was conducted. The selected parameters from the cross-correlation matrix (XP, XF, HC, CL, NFC, and $k_{h\_0.5}$) were standardized with the center (mean) and scale (standard deviation) procedure in order to evaluate the relative influence of the parameters on the dataset. According to the PCR and based on the first two PCs, 64% of the $k_{h\_0.5}$ variation in this dataset could be explained (Fig. 2a). Although all parameters (XP, XF, HC, CL, NFC) are statistically significant in this dataset, they cannot be used simultaneously as predictors, since some of them are highly correlated with each other. CL, XF and HC are strongly positively correlated, because they are located in the same direction and are close to each other. Moreover, NFC is strongly negatively correlated with CL, XF and HC, as it is located in the opposite direction (Fig. 2a). From these four parameters, only one could be used as regressor in order to minimize co-effects among the regressors. Of course, based on the plant chemical composition it is impossible to consider two parameters as independent, since they are expressed in percentage of

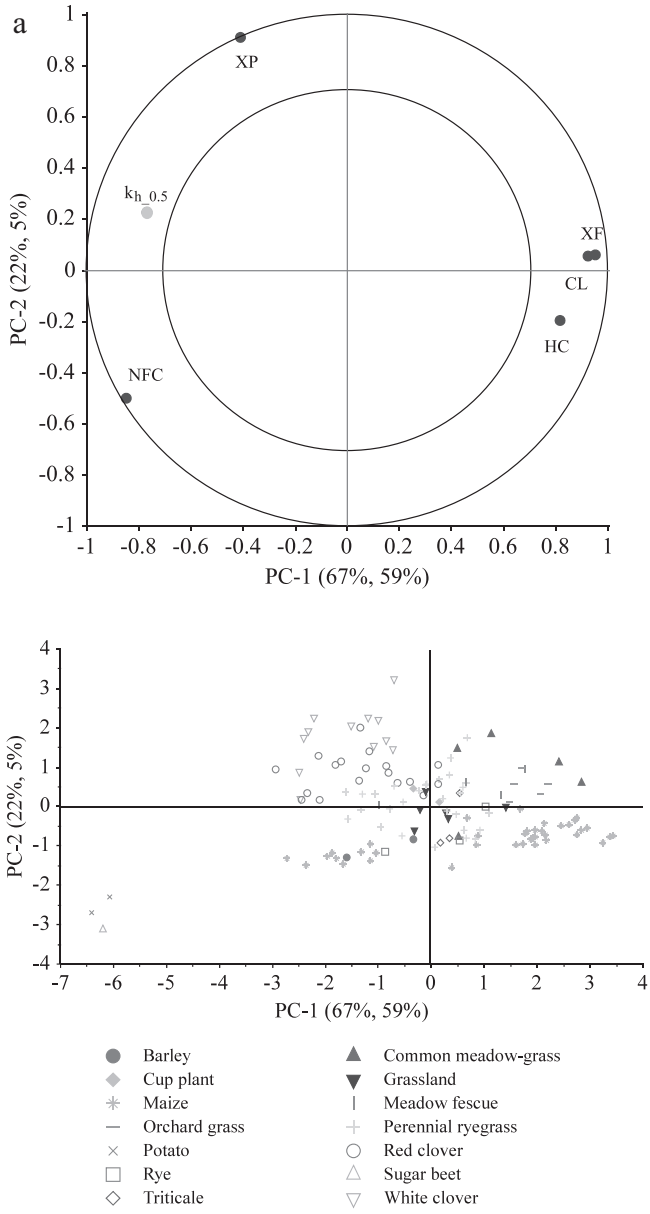

Fig. 2. a) Correlation loadings plot of principal components PC1 vs. PC2 for plant chemical compounds of the calibration dataset. The outer circle indicates 100% explained variance and the inner circle indicates 50% explained variance. b) Principal components analysis (PCA) scores for each sample.

TS [24]. NFC is closer to the outer circle (100% explanation) and moreover, it is weakly correlated with XP (almost 90° angle between them). Consequently, based on the correlation matrix and PCA, the parameters NFC and XP have been selected as regressors in the subsequent multiple linear regression analysis.

Fig. 2b exhibits the plot of the scores for each sample of the calibration's dataset. Samples that are clustered together have a similar composition. Samples with positive scores on PC1 are rich in fiber carbohydrates (CL, HC, XF) and low in NFC, such as common meadow-grass. Potato and sugar beet are far from the other samples and with negative scores on PC1. Hence, it is clear that the composition of these samples is very different. Samples with positive scores on PC2 are rich in XP, like clovers. The higher XP and NFC are, the higher the $k_{h\_0.5}$ value is and hence, the faster the biogas production process. The maize samples are clustered in two groups: one on the right side (positive score on PC1) and one on the left side of the plot (negative sore on PC1). The reason for this observation is that some of the maize samples are from the whole plant, while some are from maize stover only. This means that not only the NFC content due to present or missing kernels are different, but also the fiber content (CL, HC, XF) due to different harvesting times.

3.1.3. Mallows' Cp test

A second statistical process based on the Mallows' C_p was applied in order to define the best fitting model. The selected parameters (XP, XF, HC, CL, NFC) were the potential regressors and $k_{h\_0.5}$ was the regressand. Table 3 presents exemplary the results of the selection model method for the 5 best models. The model with the parameters NFC and XP as regressors had the lowest C_p value. Moreover, it was observed that with additional regressors both R^2 and adjusted R^2 (R^2_{adj} takes into account also the number of regressors) were not improved, and the RMSE did not change at all. Hence, additional regressors could only slightly improve the accuracy of the model calibration, while the model would become less robust for external datasets. Consequently, two different statistical approaches (PCA and Mallows' C_p) indicate NFC and XP to be the most suitable regressors for the prediction of $k_{h\_0.5}$.

3.1.4. Discussion of regressors

Potential reasons why NFC has been identified to be an important variable for the prediction of $k_{h\_0.5}$ might be as follows: Firstly, the variable NFC describes the share of all carbohydrates in the crop, since it is a percentage of VS. Secondly, carbohydrates (fiber and non-fiber) typically occupy more than 80% of the total VS of energy crops and hence, are the main source for biogas production. Thirdly, during the anaerobic process in a BMP test, the readily digestible fractions are degraded first and these are mostly characterized by the variable NFC.

XP was also statistically significant for the prediction model. Although hydrolysis rate of proteins is known to be lower than for NFC [48], the analytical parameter XP characterizes the ageing of the crop, whereas young plants are characterized by high share of XP and good anaerobic digestibility and vice versa.

3.1.5. Multiple linear regression analysis

As a next step, a multiple linear regression analysis was performed. Table 4 summarizes the results of the analysis of variance (ANOVA). Overall, the model was statistically significant with a p value as low as 0.001 and the effect of all variables (intercept, XP, NFC) were also significant (Table 4). The regression coefficient of XP was higher than the regression coefficient of NFC. However, based on the standardized regression coefficient, both regressors had similar impact on the model.

Finally, based on the calibration dataset, the following prediction model has been developed for the estimation of the hydrolysis rate constant of energy crops:

Table 3
Mallows' C_p model selection for the prediction of $k_{h\_0.5}$.

Index	C_p Value	R^2	R^2_{adj}[a]	RMSE	Regressors
1	2.4	0.662	0.657	0.00143	XP NFC
2	2.9	0.666	0.658	0.00143	XP NFC XF
3	3.7	0.669	0.658	0.00143	XP ADL NFC XF
4	3.9	0.663	0.655	0.00143	XP CL NFC
5	4.1	0.668	0.657	0.00143	XP XL NFC XF

[a] R^2_{adj}: Adjusted coefficient of determination (R^2).

Table 4
ANOVA from the multiple regression analysis.

Variable	Regression coefficients	Standard error	t value	p value	Standardized regression coefficients
Intercept	0.002	0.0004	5.33	<0.001	0
XP	0.022	0.0019	11.44	<0.001	0.59
NFC	0.012	0.0010	12.18	<0.001	0.63

$$k_{h\_0.5} = 0.002 + 0.022\,XP + 0.012\,NFC \quad (4)$$

where $k_{h\_0.5}$ is expressed in h^{-1}, and the chemical compounds are expressed in kg/kg_{VS}.

The residuals of the regression model have been tested with the Kolmogorov-Smirnov test for normal distribution at a significance level of 5% and the assumption of normality cannot be rejected. Hence, the model can be considered as valid and fairly precise.

3.2. Validation with external data

A total of 91 samples were analyzed for the validation of the model (see Table SM-2 in SM). Those samples were not included in the calibration of the model. Both Y_B and $k_{h\_0.5}$ prediction models were validated with this dataset. Table 5 summarizes the results of the validation. Both models predicted the reference values with high accuracy and explained bulk of variation in the dataset. However, the models' performances differed. Potential reasons are discussed below.

3.2.1. Assessment of model performance for the prediction of Y_B

The biogas yield model could predict measured values with an average prediction error (CVRMSE) of 10%. Although the mean of the predicted values was very close to the mean of the measured values, the range of the values could not be predicted so well (Table 5). The validation of the Y_B model has shown that the extreme values of the dataset (min and max) were predicted with low accuracy. However, the model could predict the dataset with a coefficient of variation (CV) similar to the measured data and with a remarkable correlation coefficient of 0.62 for such a diverse dataset. The RMSE is relatively high for the Y_B model. However, this is a typical characteristic of global models, as they are suitable for diverse substrates, but with lower accuracy than plant species specific models.

3.2.2. Assessment of model performance for the prediction of $k_{h\_0.5}$

The model for $k_{h\_0.5}$ performed worse compared to the Y_B model with an average prediction error (CVRMSE) of 15%. The range of the dataset was underestimated and the CV was 10 percentage points lower than the one of the measurements. However, the mean of the predicted values was similar to the mean of the measured values, and the predicted values were strongly correlated with the measured values as indicated by a correlation coefficient of 0.93 (Table 5). The results indicate that the $k_{h\_0.5}$ model could predict the differentiation in the dataset very well. This is further manifested in the plot of measured versus predicted values, as the points were very close to their regression line with a R^2 value of 0.86 (Fig. 3).

Results presented in Fig. 3 show that the model performed better for the range of $k_{h\_0.5}$ between 0.004 and 0.009. For measurements above 0.009, the model underestimated the values. Consequently, the CVRMSE was relatively high (15%). The model has particularly shown poor performance for the samples of "virginia mallow". Although this plant species was characterized by high XP and NFC values, the model clearly underestimated $k_{h\_0.5}$. Since the $k_{h\_0.5}$ value of "virginia mallow" plants was within the calibration range of $k_{h\_0.5}$, this global model was able to predict the differentiation among the samples. However, due to fact that the plant species "virginia mallow" has not been included in the calibration, the model could only poorly predict the $k_{h\_0.5}$ value. This moderate prediction accuracy might also be explained by the fact that the complex microbial processes are approximated by a simple first-order kinetic model. It has been shown that more complex kinetic models can simulate more accurately the biogas production than a first-order model [6]. However, the simplicity of a first-order model makes this approach very attractive to rapid feedstock assessment.

Table 5
Parameters of the independent validation of biogas yield prediction model (Eq. (2)) and hydrolysis rate constant prediction model (Eq. (4)).

Parameter	Y_B measured	Y_B Eq. (2)	$k_{h\_0.5}$ measured	$k_{h\_0.5}$ Eq. (4)
Min	378	410	0.0040	0.0045
Max	720	678	0.0151	0.0113
Range	342	268	0.0111	0.0068
Mean	576	561	0.0069	0.0066
CV	11%	12%	31%	21%
RMSE	–	60	–	0.0011
CVRMSE	–	10%	–	15%
r	–	0.62	–	0.93

3.2.3. Assessment of model performance for the prediction of BMP test

Two samples were selected to visualize the models' performance. Based on the best and worst performance of the model, sample V-4 (cup plant) and sample V-59 (virginia mallow) from the validation dataset were selected (Table 6).

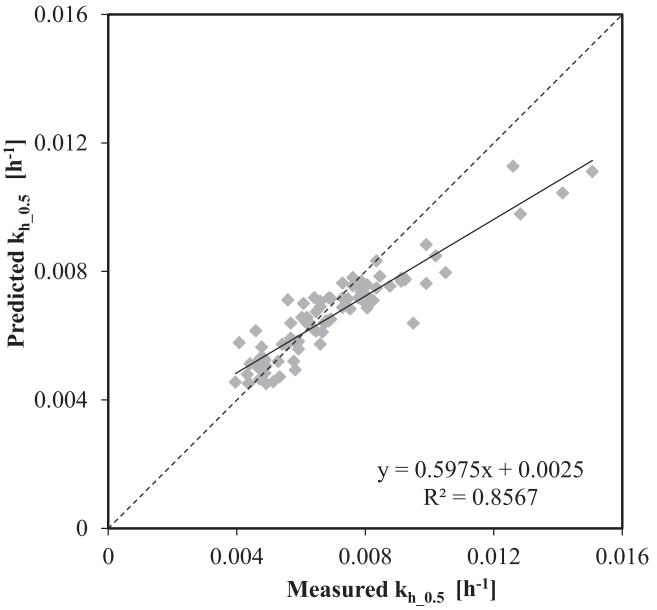

Fig. 3. Measured vs. predicted values of hydrolysis rate constant ($k_{h\_0.5}$) from the validation of the model (Eq. (4)).

Table 6
XP, HC, ADL and NFC content of the two selected samples.

Sample Nr.	Common Name	XP [kg/kg$_{VS}$]	HC [kg/kg$_{VS}$]	ADL [kg/kg$_{VS}$]	NFC [kg/kg$_{VS}$]
V-4	Cup plant	0.076	0.153	0.077	0.209
V-59	Virginia mallow	0.168	0.103	0.045	0.450

Fig. 4 shows the measured and the predicted values for the whole duration of the trial, whereas biogas yields were predicted according to Eq. (2) (Y_B model), and biogas production rates were predicted according to Eq. (1) (first-order kinetic model) based on the $k_{h\_0.5}$ model (Eq. (4)). As for sample V-4, Y_B was well predicted and the biogas production was also precisely estimated until the half of Y_B. Between the half and the final value of Y_B, the model slightly overestimated the biogas production (Fig. 4).

As for sample V-59, the biogas production was underestimated until the half of Y_B, after which the biogas production and the Y_B were overestimated. In order to explain the different behavior of the models, the chemical compositions of the two samples have to be considered. About 33% lower HC and 42% lower ADL was recorded for sample V-59 than for sample V-4 (Table 6) and since ADL has a higher impact on the Y_B model, the value was overestimated. Additionally, more than the double amount of XP and NFC was recorded for sample V-59 and, as already mentioned, the plant species of sample V-59 was not included in the dataset of the model calibration. Hence, the global model could only roughly predict the hourly biogas production of the unknown species. Being able to at least roughly predict the BMP curve of an unknown species, the model has proven its suitability as a global model. By extending the calibration to more species, the existing model can be modified in future studies.

The biochemical processes of anaerobic digestion are very complex and cannot be precisely described by a simple first-order kinetic model as has been concluded by Li et al. [49] already. However, both models could accurately predict the differentiation of the samples and these results depict the main advantage of this study's approach.

Based on the model's prediction, the biogas yield of sample V-59 was 25% higher than of sample V-4 and the time to reach the half-maximum biogas production of sample V-59 was 54% less than of sample V-4. Based on the experimental data, the biogas yield of sample V-59 was 19% higher than of sample V-4 and the time to reach the half-maximum biogas production of sample V-59 was 43% less than of sample V-4. Although the values could not be estimated with high accuracy, both results lead to the same conclusion concerning the feedstock characterization. Hence, feedstock ranking with this approach is possible, which was the main aim of the study.

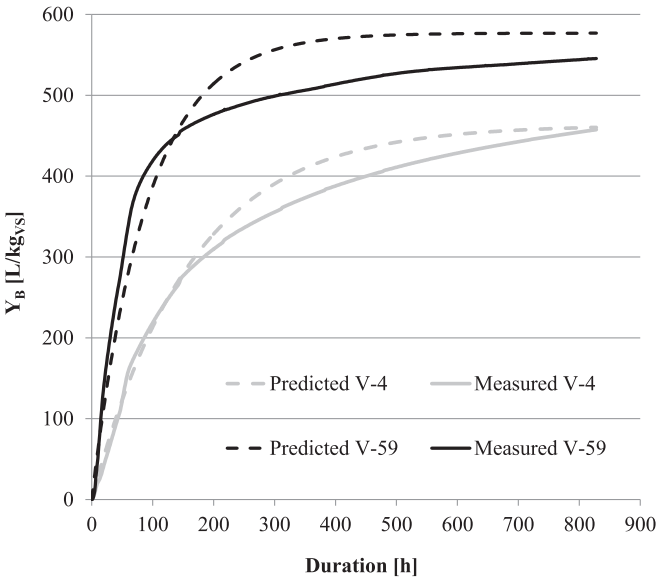

Fig. 4. Cumulative biogas production in BMP test (solid curve) and the prediction (dashed curve) for cup plant (V-4) and virginia mallow (V-59).

The model is able to predict the different biogas production rates of diverse samples and it can be applied for substrate ranking. Consequently, linear regression models can be utilized for substrate assessment, defining differences between samples. However, the prediction accuracy of individual samples may be less satisfying, since systematic effects (e.g., microclimate of local cultivation conditions, new cultivars, etc.) cannot be predicted by such static models. Furthermore, the prediction of the biogas production rate can provide additional information about the determination of an optimal phenological stage of plants in order to optimize the feedstock's quality for agricultural biogas plants providing biogas on-demand.

3.3. Outlook and future research

The first-order kinetic model depicted the differentiation of the samples well. The proposed approach for biogas production in BMP test prediction was developed based on large datasets of various energy crops. Hence, the model can be considered as a global model enabling the rough estimation even of unknown species. Future studies should enhance the existing model by extending the calibration dataset with additional plant species or cultivars. Specific models can be developed following the proposed approach in order to identify variables that could minimize the estimation error.

4. Conclusion

The results of this study indicate that a first-order model can reproduce the biogas production rate in a BMP test and linear regression models can precisely predict the differentiation in the biogas production of different energy crops. Despite the fact that the prediction is limited due to its simplicity and cannot accommodate extreme cases, this approach can be a useful tool for practitioners in order to assess the suitability for biogas production of different feedstocks. This is a first step towards the development of a global model with high precision for the prediction of both biogas yield and production rate simultaneously.

Acknowledgement

The authors thank Dr. Stephan Hartmann, Dorothea Hofmann, Dr. Joachim Eder, Monika Fleschhut, Dr. Fritz und Dr. Anja Hartmann for providing the samples and for their support. Anke Aschmann, Natascha Siddiqui and Michael Hanrieder are kindly acknowledged for their excellent technical assistance. This project was funded by the Bavarian State Ministry of Food, Agriculture and Forestry (BE/15/06) and the Bavarian State Ministry of Economic Affairs and Media, Energy and Technology, Project code: BE/15/06.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.renene.2017.10.100.

References

[1] V. Uusitalo, R. Soukka, M. Horttanainen, A. Niskanen, J. Havukainen, Economics and greenhouse gas balance of biogas use systems in the finnish transportation sector, Renew. Energy 51 (2013) 132–140, https://doi.org/10.1016/j.renene.2012.09.002.

[2] A. Meyer-Aurich, A. Schattauer, H.J. Hellebrand, H. Klauss, M. Plöchl, W. Berg, Impact of uncertainties on greenhouse gas mitigation potential of biogas production from agricultural resources, Renew. Energy 37 (2012) 277–284, https://doi.org/10.1016/j.renene.2011.06.030.

[3] M. Esen, T. Yuksel, Experimental evaluation of using various renewable energy sources for heating a greenhouse, Energy Build. 65 (2013) 340–351, https://doi.org/10.1016/j.enbuild.2013.06.018.

[4] O. Ellabban, H. Abu-Rub, F. Blaabjerg, Renewable energy resources: current status, future prospects and their enabling technology, Renew. Sustain. Energy Rev. 39 (2014) 748–764, https://doi.org/10.1016/j.rser.2014.07.113.

[5] S. Auburger, A. Jacobs, B. Märländer, E. Bahrs, Economic optimization of feedstock mix for energy production with biogas technology in Germany with a special focus on sugar beets – effects on greenhouse gas emissions and energy balances, Renew. Energy 89 (2016) 1–11, https://doi.org/10.1016/j.renene.2015.11.042.

[6] P. Tsapekos, P.G. Kougias, H. Egelund, U. Larsen, J. Pedersen, P. Trénel, I. Angelidaki, Mechanical pretreatment at harvesting increases the bioenergy output from marginal land grasses, Renew. Energy 111 (2017) 914–921, https://doi.org/10.1016/j.renene.2017.04.061.

[7] H. Lund, A.N. Andersen, P.A. Østergaard, B.V. Mathiesen, D. Connolly, From electricity smart grids to smart energy systems – a market operation based approach and understanding, Energy 42 (2012) 96–102, https://doi.org/10.1016/j.energy.2012.04.003.

[8] F. Steinke, P. Wolfrum, C. Hoffmann, Grid vs. storage in a 100% renewable Europe, Renew. Energy 50 (2013) 826–832, https://doi.org/10.1016/j.renene.2012.07.044.

[9] E. Mauky, S. Weinrich, H.-J. Nägele, H.F. Jacobi, J. Liebetrau, M. Nelles, Model predictive control for demand-driven biogas production in full scale, Chem. Eng. Technol. 39 (2016) 652–664, https://doi.org/10.1002/ceat.201500412.

[10] A. Lemmer, J. Krümpel, Demand-driven biogas production in anaerobic filters, Appl. Energy 185 (Part 1) (2017) 885–894, https://doi.org/10.1016/j.apenergy.2016.10.073.

[11] H. Hahn, W. Ganagin, K. Hartmann, M. Wachendorf, Cost analysis of concepts for a demand oriented biogas supply for flexible power generation, Bioresour. Technol. 170 (2014) 211–220, https://doi.org/10.1016/j.biortech.2014.07.085.

[12] C. Mao, Y. Feng, X. Wang, G. Ren, Review on research achievements of biogas from anaerobic digestion, Renew. Sustain. Energy Rev. 45 (2015) 540–555, https://doi.org/10.1016/j.rser.2015.02.032.

[13] S. Weinrich, M. Nelles, Critical comparison of different model structures for the applied simulation of the anaerobic digestion of agricultural energy crops, Bioresour. Technol. 178 (2015) 306–312, https://doi.org/10.1016/j.biortech.2014.10.138.

[14] S.G. Pavlostathis, E. Giraldo-Gomez, Kinetics of anaerobic treatment: a critical review, Crit. Rev. Environ. Control 21 (1991) 411–490, https://doi.org/10.1080/10643389109388424.

[15] K. Koch, J.E. Drewes, Alternative approach to estimate the hydrolysis rate constant of particulate material from batch data, Appl. Energy 120 (2014) 11–15, https://doi.org/10.1016/j.apenergy.2014.01.050.

[16] I. Angelidaki, M. Alves, D. Bolzonella, L. Borzacconi, J.L. Campos, A.J. Guwy, S. Kalyuzhnyi, P. Jenicek, J.B. van Lier, Defining the biomethane potential (BMP) of solid organic wastes and energy crops: a proposed protocol for batch assays, Water Sci. Technol. 59 (2009) 927–934, https://doi.org/10.2166/wst.2009.040.

[17] VDI, Guideline VDI 4630-Fermentation of Organic Materials. Characterisation of the Substrate, Sampling, Collection of Material Data, Fermentation Tests, Beuth-Verlang, Düsseldorf, Germany, 2006.

[18] S. Wulf, K. Tillman, M. Bischoff, Measurement of Biogas and Methane Yields in Fermentation Tests, VDLUFA-Verlang, Darmstadt, Germany, 2013.

[19] F. Raposo, V. Fernández-Cegrí, M. a De la Rubia, R. Borja, F. Béline, C. Cavinato, G. Demirer, B. Fernández, M. Fernández-Polanco, J. c. Frigon, R. Ganesh, P. Kaparaju, J. Koubova, R. Méndez, G. Menin, A. Peene, P. Scherer, M. Torrijos, H. Uellendahl, I. Wierinck, V. de Wilde, Biochemical methane potential (BMP) of solid organic substrates: evaluation of anaerobic biodegradability using data from an international interlaboratory study, J. Chem. Technol. Biotechnol. 86 (2011) 1088–1098, https://doi.org/10.1002/jctb.2622.

[20] C. Holliger, M. Alves, D. Andrade, I. Angelidaki, S. Astals, U. Baier, C. Bougrier, P. Buffière, M. Carbella, V. de Wilde, F. Ebertseder, B. Fernández, E. Ficara, I. Fotidis, J.-C. Frigon, H. Fruteau de Laclos, D.S.M. Ghasimi, G. Hack, M. Hartel, J. Heerenklage, I. Sarvari Horvath, P. Jenicek, K. Koch, J. Krautwald, J. Lizasoain, J. Liu, L. Mosberger, M. Nistor, H. Oechsner, J.V. Oliveira, M. Paterson, A. Pauss, S. Pommier, I. Porqueddu, F. Raposo, T. Ribeiro, F. Rüsch Pfund, S. Strömberg, M. Torrijos, M. van Eekert, J. van Lier, H. Wedwitschka, I. Wierinck, Towards a standardization of biomethane potential tests, Water Sci. Technol. 74 (2016). https://infoscience.epfl.ch/record/219009. (Accessed 19 July 2016).

[21] Q. Zhang, J. Hu, D.-J. Lee, Biogas from anaerobic digestion processes: research updates, Renew. Energy 98 (2016) 108–119, https://doi.org/10.1016/j.renene.2016.02.029.

[22] V. Dandikas, H. Heuwinkel, F. Lichti, J.E. Drewes, K. Koch, Correlation between biogas yield and chemical composition of energy crops, Bioresour. Technol. 174 (2014) 316–320, https://doi.org/10.1016/j.biortech.2014.10.019.

[23] J.M. Triolo, S.G. Sommer, H.B. Møller, M.R. Weisbjerg, X.Y. Jiang, A new algorithm to characterize biodegradability of biomass during anaerobic digestion: influence of lignin concentration on methane production potential, Bioresour. Technol. 102 (2011) 9395–9402, https://doi.org/10.1016/j.biortech.2011.07.026.

[24] S.T. Thomsen, H. Spliid, H. Østergård, Statistical prediction of biomethane potentials based on the composition of lignocellulosic biomass, Bioresour. Technol. 154 (2014) 80–86, https://doi.org/10.1016/j.biortech.2013.12.029.

[25] S. Strömberg, M. Nistor, J. Liu, Early prediction of Biochemical Methane Potential through statistical and kinetic modelling of initial gas production, Bioresour. Technol. 176 (2015) 233–241, https://doi.org/10.1016/j.biortech.2014.11.033.

[26] R. Mulka, W. Szulczewski, J. Szlachta, M. Mulka, Estimation of methane production for batch technology – a new approach, Renew. Energy 90 (2016) 440–449, https://doi.org/10.1016/j.renene.2016.01.012.

[27] D.J. Batstone, D. Puyol, X. Flores-Alsina, J. Rodríguez, Mathematical modelling of anaerobic digestion processes: applications and future needs, Rev. Environ. Sci. Biotechnol. 14 (2015) 595–613, https://doi.org/10.1007/s11157-015-9376-4.

[28] L. Blasco, M. Kahala, E. Tampio, S. Ervasti, T. Paavola, J. Rintala, V. Joutsjoki, Dynamics of microbial communities in untreated and autoclaved food waste anaerobic digesters, Anaerobe 29 (2014) 3–9, https://doi.org/10.1016/j.anaerobe.2014.04.011.

[29] M. Lebuhn, A. Hanreich, M. Klocke, A. Schlüter, C. Bauer, C.M. Pérez, Towards molecular biomarkers for biogas production from lignocellulose-rich substrates, Anaerobe 29 (2014) 10–21, https://doi.org/10.1016/j.anaerobe.2014.04.006.

[30] G. Willeghems, J. Buysse, Changing old habits: the case of feeding patterns in anaerobic digesters, Renew. Energy 92 (2016) 212–221, https://doi.org/10.1016/j.renene.2016.01.081.

[31] A.G. Hashimoto, Effect of inoculum/substrate ratio on methane yield and production rate from straw, Biol. Wastes 28 (1989) 247–255, https://doi.org/10.1016/0269-7483(89)90108-0.

[32] G. Bekiaris, J.M. Triolo, C. Peltre, L. Pedersen, L.S. Jensen, S. Bruun, Rapid estimation of the biochemical methane potential of plant biomasses using Fourier transform mid-infrared photoacoustic spectroscopy, Bioresour. Technol. 197 (2015) 475–481, https://doi.org/10.1016/j.biortech.2015.08.050.

[33] T. Edwiges, L. Frare, B. Mayer, L. Lins, J.M. Triolo, X. Flotats, M.S.S. de Mendonça Costa, Influence of chemical composition on biochemical methane potential of fruit and vegetable waste, Waste Manag. (2017), https://doi.org/10.1016/j.wasman.2017.05.030.

[34] J.M. Triolo, A.J. Ward, L. Pedersen, M.M. Løkke, H. Qu, S.G. Sommer, Near Infrared Reflectance Spectroscopy (NIRS) for rapid determination of biochemical methane potential of plant biomass, Appl. Energy 116 (2014) 52–57, https://doi.org/10.1016/j.apenergy.2013.11.006.

[35] J. De Vrieze, L. Raport, B. Willems, S. Verbrugge, E. Volcke, E. Meers, L.T. Angenent, N. Boon, Inoculum selection influences the biochemical methane potential of agro-industrial substrates, Microb. Biotechnol. 8 (2015) 776–786, https://doi.org/10.1111/1751-7915.12268.

[36] Y. Gu, X. Chen, Z. Liu, X. Zhou, Y. Zhang, Effect of inoculum sources on the anaerobic digestion of rice straw, Bioresour. Technol. 158 (2014) 149–155, https://doi.org/10.1016/j.biortech.2014.02.011.

[37] J.D. Browne, J.D. Murphy, Assessment of the resource associated with biomethane from food waste, Appl. Energy 104 (2013) 170–177, https://doi.org/10.1016/j.apenergy.2012.11.017.

[38] K. Koch, T. Lippert, J.E. Drewes, The role of inoculum's origin on the methane yield of different substrates in biochemical methane potential (BMP) tests, Bioresour. Technol. 243 (2017) 457–463, https://doi.org/10.1016/j.biortech.2017.06.142.

[39] V. Dandikas, H. Heuwinkel, F. Lichti, J.E. Drewes, K. Koch, Correlation between biogas yield and chemical composition of grassland plant species, Energy fuels. 29 (2015) 7221–7229, https://doi.org/10.1021/acs.energyfuels.5b01257.

[40] DEV, Deutsche Einheitsverfahren zur Wasser-, Abwasser- und Schlammuntersuchung (German standard Methods for the Examination of Water, Wastewater and Sludge), VCH, Wiley, 2015.

[41] P.J. Van Soest, R.H. Wine, Use of detergents in the analysis of fibrous feeds. IV. Determination of plant cell-wall constituents, J. Assoc. Off. Agric. Chem. 50 (1967) 50–55.

[42] K. Naumann, R. Bassler, Chemical Analysis of Feedstuffs, VDLUFA-Verlang, Darmstadt, Germany, 1976.

[43] B.G. Tabachnick, L.S. Fidell, Using Multivariate Statistics, Pearson Education, Boston, 2012.

[44] H.N. Gavala, I. Angelidaki, B.K. Ahring, Kinetics and modeling of anaerobic digestion process, in: P.B.K. Ahring, I. Angelidaki, E.C. de Macario, H.N. Gavala, J. Hofman-Bang, A.J.L. Macario, S.J.W.H.O. Elferink, L. Raskin, A.J.M. Stams, P. Westermann, D. Zheng (Eds.), Biomethanation I, Springer Berlin Heidelberg, 2003, pp. 57–93, https://doi.org/10.1007/3-540-45839-5_3.

[45] C. Herrmann, C. Idler, M. Heiermann, Biogas crops grown in energy crop rotations: linking chemical composition and methane production characteristics, Bioresour. Technol. 206 (2016) 23–35, https://doi.org/10.1016/j.biortech.2016.01.058.
[46] J. Rath, H. Heuwinkel, A. Herrmann, Specific biogas yield of maize can Be predicted by the interaction of four biochemical constituents, BioEnergy Res. (2013) 1–14, https://doi.org/10.1007/s12155-013-9318-3.
[47] K. Koch, M. Luebken, T. Gehring, M. Wichern, H. Horn, Biogas from grass silage - measurements and modeling with ADM1, Bioresour. Technol. 101 (2010) 8158–8165, https://doi.org/10.1016/j.biortech.2010.06.009.
[48] M. Lübken, K. Koch, T. Gehring, H. Horn, M. Wichern, Parameter estimation and long-term process simulation of a biogas reactor operated under trace elements limitation, Appl. Energy 142 (2015) 352–360, https://doi.org/10.1016/j.apenergy.2015.01.014.
[49] D. Li, X. Huang, Q. Wang, Y. Yuan, Z. Yan, Z. Li, Y. Huang, X. Liu, Kinetics of methane production and hydrolysis in anaerobic digestion of corn stover, Energy 102 (2016) 1–9, https://doi.org/10.1016/j.energy.2016.02.074.

Correlation between hydrolysis rate constant and chemical composition of energy crops

Vasilis Dandikas[a], Hauke Heuwinkel[b], Fabian Lichti[a], Thomas Eckl[c], Jörg E. Drewes[d] and Konrad Koch[d*]

[a] Institute for Agricultural Engineering and Animal Husbandry, Bavarian State Research Center for Agriculture, Am Staudengarten 3, 85354 Freising, Germany.

[b] Department of Agriculture and Food Economy, Hochschule Weihenstephan-Triesdorf, Am Staudengarten 1, 85354 Freising, Germany.

[c] Department of Agricultural Trial Operation, Bavarian State Research Center for Agriculture, Lange Point 12, 85354 Freising, Germany

[d] Chair of Urban Water Systems Engineering, Technical University of Munich, Am Coulombwall 3, 85748 Garching, Germany.

- Supplementary Material -

[*] Corresponding author: Konrad Koch. E-mail address: k.koch@tum.de

Table SM-1: Chemical composition in kg/kg_{VS}, biogas yield in L/kg_{VS}, and hydrolysis rate constant in h^{-1} of dataset's energy crops for model calibration.

Nr.	Common Name	XP	XL	XF	HC	CL	ADL	NFC	Y_B	$k_{h\_0.5}$
C-1	Barley	0.082	0.022	0.236	0.233	0.272	0.026	0.365	663	0.0094
C-2	Barley	0.076	0.018	0.184	0.220	0.215	0.019	0.452	617	0.0108
C-3	Common meadow-grass	0.120	0.013	0.319	0.401	0.358	0.039	0.068	488	0.0046
C-4	Common meadow-grass	0.070	0.017	0.229	0.312	0.293	0.036	0.272	717	0.0061
C-5	Common meadow-grass	0.155	0.026	0.328	0.343	0.353	0.043	0.080	642	0.0071
C-6	Common meadow-grass	0.219	0.031	0.273	0.333	0.290	0.029	0.097	682	0.0090
C-7	Common meadow-grass	0.198	0.031	0.240	0.265	0.306	0.048	0.152	704	0.0091
C-8	Cup plant	0.102	0.050	0.270	0.060	0.398	0.075	0.316	443	0.0071
C-9	Cup plant	0.136	0.045	0.225	0.122	0.347	0.063	0.288	446	0.0080
C-10	Grassland	0.083	0.033	0.355	0.104	0.412	0.090	0.279	431	0.0086
C-11	Grassland	0.149	0.032	0.240	0.267	0.255	0.043	0.254	653	0.0089
C-12	Grassland	0.100	0.014	0.249	0.265	0.291	0.053	0.277	555	0.0043
C-13	Grassland	0.125	0.031	0.230	0.257	0.268	0.021	0.297	715	0.0078
C-14	Grassland	0.094	0.022	0.243	0.212	0.309	0.116	0.247	356	0.0038
C-15	Grassland	0.077	0.049	0.256	0.121	0.326	0.063	0.364	490	0.0092
C-16	Maize	0.037	0.010	0.301	0.304	0.377	0.028	0.245	596	0.0062
C-17	Maize	0.058	0.008	0.320	0.322	0.391	0.029	0.193	576	0.0055
C-18	Maize	0.044	0.005	0.324	0.345	0.405	0.034	0.167	595	0.0049
C-19	Maize	0.025	0.005	0.332	0.425	0.371	0.042	0.131	465	0.0030
C-20	Maize	0.030	0.005	0.327	0.435	0.379	0.056	0.096	453	0.0032
C-21	Maize	0.025	0.005	0.343	0.405	0.377	0.076	0.112	488	0.0035
C-22	Maize	0.040	0.006	0.330	0.237	0.371	0.042	0.303	592	0.0058
C-23	Maize	0.041	0.006	0.317	0.321	0.368	0.041	0.224	651	0.0056

C-24	Maize	0.068	0.005	0.308	0.374	0.389	0.025	0.139	604	0.0058
C-25	Maize	0.052	0.007	0.324	0.272	0.391	0.040	0.238	587	0.0055
C-26	Maize	0.044	0.007	0.306	0.281	0.382	0.042	0.244	606	0.0058
C-27	Maize	0.057	0.008	0.346	0.280	0.421	0.049	0.185	599	0.0056
C-28	Maize	0.050	0.008	0.311	0.344	0.352	0.039	0.206	636	0.0059
C-29	Maize	0.070	0.006	0.293	0.357	0.360	0.023	0.183	624	0.0066
C-30	Maize	0.024	0.005	0.278	0.210	0.324	0.045	0.391	637	0.0069
C-31	Maize	0.046	0.008	0.343	0.350	0.388	0.052	0.155	599	0.0062
C-32	Maize	0.048	0.007	0.305	0.311	0.358	0.037	0.239	618	0.0069
C-33	Maize	0.041	0.006	0.347	0.337	0.383	0.043	0.191	584	0.0062
C-34	Maize	0.051	0.008	0.316	0.319	0.350	0.039	0.233	581	0.0061
C-35	Maize	0.058	0.006	0.356	0.302	0.389	0.044	0.200	610	0.0065
C-36	Maize	0.055	0.007	0.310	0.258	0.389	0.040	0.251	629	0.0070
C-37	Maize	0.057	0.007	0.304	0.304	0.356	0.033	0.242	658	0.0070
C-38	Maize	0.042	0.007	0.328	0.256	0.369	0.039	0.287	626	0.0071
C-39	Maize	0.058	0.007	0.320	0.291	0.370	0.039	0.236	574	0.0071
C-40	Maize	0.074	0.021	0.217	0.226	0.233	0.025	0.421	673	0.0124
C-41	Maize	0.066	0.023	0.204	0.253	0.215	0.017	0.428	680	0.0124
C-42	Maize	0.067	0.025	0.184	0.221	0.207	0.021	0.459	646	0.0126
C-43	Maize	0.075	0.029	0.155	0.203	0.182	0.013	0.498	685	0.0126
C-44	Maize	0.073	0.023	0.171	0.200	0.222	0.023	0.459	715	0.0128
C-45	Maize	0.079	0.038	0.183	0.154	0.234	0.025	0.471	733	0.0090
C-46	Maize	0.088	0.035	0.138	0.175	0.182	0.012	0.507	799	0.0092
C-47	Maize	0.073	0.033	0.204	0.189	0.239	0.034	0.432	706	0.0096
C-48	Maize	0.080	0.035	0.159	0.215	0.201	0.019	0.450	640	0.0096
C-49	Maize	0.054	0.012	0.264	0.280	0.326	0.026	0.302	617	0.0081
C-50	Maize	0.068	0.012	0.266	0.279	0.325	0.027	0.289	581	0.0073
C-51	Meadow fescue	0.157	0.013	0.251	0.305	0.298	0.022	0.206	545	0.0071

C-52	Meadow fescue	0.121	0.022	0.293	0.272	0.347	0.024	0.215	687	0.0078
C-53	Meadow fescue	0.153	0.024	0.330	0.266	0.355	0.038	0.164	646	0.0081
C-54	Meadow fescue	0.139	0.033	0.197	0.222	0.251	0.024	0.332	621	0.0095
C-55	Orchard grass	0.109	0.037	0.327	0.304	0.352	0.039	0.158	536	0.0054
C-56	Orchard grass	0.108	0.036	0.301	0.305	0.321	0.037	0.194	570	0.0062
C-57	Orchard grass	0.123	0.025	0.335	0.298	0.366	0.040	0.149	646	0.0076
C-58	Orchard grass	0.132	0.027	0.291	0.283	0.354	0.034	0.169	679	0.0079
C-59	Orchard grass	0.152	0.030	0.314	0.284	0.339	0.059	0.136	633	0.0090
C-60	Perennial ryegrass	0.144	0.014	0.213	0.272	0.246	0.022	0.301	572	0.0068
C-61	Perennial ryegrass	0.105	0.021	0.262	0.251	0.311	0.027	0.285	671	0.0085
C-62	Perennial ryegrass	0.136	0.023	0.280	0.253	0.298	0.041	0.250	748	0.0087
C-63	Perennial ryegrass	0.163	0.038	0.214	0.260	0.219	0.061	0.259	644	0.0090
C-64	Perennial ryegrass	0.153	0.027	0.236	0.247	0.269	0.023	0.282	768	0.0090
C-65	Perennial ryegrass	0.078	0.023	0.227	0.231	0.281	0.030	0.357	639	0.0091
C-66	Perennial ryegrass	0.138	0.030	0.189	0.213	0.235	0.026	0.358	676	0.0107
C-67	Perennial ryegrass	0.108	0.030	0.205	0.226	0.247	0.021	0.368	637	0.0108
C-68	Perennial ryegrass	0.163	0.034	0.231	0.291	0.271	0.025	0.216	597	0.0062
C-69	Perennial ryegrass	0.220	0.036	0.245	0.319	0.286	0.018	0.121	584	0.0073
C-70	Perennial ryegrass	0.078	0.019	0.255	0.276	0.308	0.039	0.281	609	0.0071
C-71	Perennial ryegrass	0.114	0.020	0.258	0.257	0.294	0.028	0.287	675	0.0084
C-72	Perennial ryegrass	0.131	0.026	0.240	0.254	0.298	0.035	0.257	669	0.0088
C-73	Perennial ryegrass	0.165	0.029	0.259	0.251	0.292	0.042	0.221	732	0.0096
C-74	Perennial ryegrass	0.146	0.026	0.273	0.267	0.298	0.028	0.235	715	0.0096
C-75	Perennial ryegrass	0.152	0.039	0.192	0.197	0.261	0.036	0.315	655	0.0103
C-76	Perennial ryegrass	0.192	0.040	0.243	0.289	0.274	0.037	0.168	612	0.0062
C-77	Perennial ryegrass	0.076	0.017	0.289	0.263	0.323	0.028	0.294	630	0.0071
C-78	Perennial ryegrass	0.062	0.016	0.294	0.244	0.322	0.038	0.317	598	0.0083
C-79	Perennial ryegrass	0.090	0.024	0.214	0.243	0.265	0.016	0.361	662	0.0085

C-80	Perennial ryegrass	0.062	0.014	0.263	0.228	0.296	0.028	0.371	625	0.0085
C-81	Perennial ryegrass	0.160	0.031	0.240	0.249	0.274	0.022	0.264	691	0.0095
C-82	Perennial ryegrass	0.130	0.032	0.179	0.206	0.228	0.014	0.391	663	0.0098
C-83	Perennial ryegrass	0.169	0.037	0.176	0.194	0.231	0.022	0.346	663	0.0114
C-84	Perennial ryegrass	0.097	0.033	0.265	0.310	0.315	0.025	0.220	594	0.0067
C-85	Perennial ryegrass	0.158	0.013	0.247	0.343	0.270	0.025	0.192	549	0.0067
C-86	Perennial ryegrass	0.127	0.024	0.226	0.166	0.288	0.037	0.357	652	0.0078
C-87	Perennial ryegrass	0.069	0.015	0.280	0.260	0.308	0.032	0.316	631	0.0083
C-88	Perennial ryegrass	0.161	0.035	0.185	0.207	0.242	0.025	0.330	629	0.0087
C-89	Perennial ryegrass	0.159	0.031	0.234	0.258	0.274	0.026	0.252	699	0.0096
C-90	Potato	0.087	0.003	0.047	0.040	0.074	0.001	0.795	746	0.0147
C-91	Potato	0.070	0.004	0.040	0.027	0.061	0.001	0.839	700	0.0133
C-92	Red clover	0.140	0.033	0.242	0.218	0.293	0.027	0.289	619	0.0085
C-93	Red clover	0.167	0.038	0.217	0.210	0.273	0.021	0.291	713	0.0083
C-94	Red clover	0.221	0.028	0.146	0.131	0.180	0.046	0.394	531	0.0094
C-95	Red clover	0.141	0.028	0.295	0.104	0.350	0.073	0.303	519	0.0100
C-96	Red clover	0.162	0.021	0.185	0.064	0.245	0.051	0.456	577	0.0107
C-97	Red clover	0.157	0.019	0.198	0.076	0.256	0.055	0.437	540	0.0110
C-98	Red clover	0.157	0.020	0.273	0.110	0.320	0.066	0.328	547	0.0110
C-99	Red clover	0.225	0.021	0.215	0.060	0.251	0.057	0.385	629	0.0110
C-100	Red clover	0.209	0.025	0.196	0.147	0.225	0.059	0.336	539	0.0112
C-101	Red clover	0.178	0.021	0.219	0.137	0.258	0.051	0.354	537	0.0112
C-102	Red clover	0.174	0.022	0.195	0.088	0.228	0.052	0.435	524	0.0114
C-103	Red clover	0.178	0.024	0.241	0.129	0.293	0.064	0.313	535	0.0116
C-104	Red clover	0.183	0.019	0.201	0.130	0.317	0.073	0.278	515	0.0122
C-105	Red clover	0.217	0.027	0.215	0.149	0.262	0.066	0.279	596	0.0124
C-106	Red clover	0.172	0.020	0.273	0.137	0.344	0.070	0.257	510	0.0124
C-107	Red clover	0.186	0.016	0.242	0.065	0.302	0.078	0.353	553	0.0128

C-108	Red clover	0.256	0.028	0.212	0.163	0.237	0.071	0.245	581	0.0131
C-109	Red clover	0.208	0.020	0.231	0.063	0.274	0.061	0.374	555	0.0136
C-110	Rye	0.068	0.017	0.240	0.165	0.276	0.030	0.442	629	0.0071
C-111	Rye	0.064	0.016	0.283	0.221	0.323	0.033	0.343	713	0.0072
C-112	Rye	0.107	0.024	0.290	0.266	0.322	0.040	0.241	642	0.0078
C-113	Sorghum	0.090	0.027	0.279	0.223	0.324	0.062	0.274	514	0.0057
C-114	Sorghum	0.086	0.014	0.219	0.176	0.262	0.025	0.436	665	0.0057
C-115	Sorghum	0.089	0.021	0.296	0.290	0.350	0.061	0.189	464	0.0044
C-116	Sugar beet	0.041	0.001	0.041	0.058	0.057	0.003	0.840	700	0.0107
C-117	Triticale	0.138	0.021	0.285	0.234	0.311	0.027	0.269	680	0.0071
C-118	Triticale	0.063	0.025	0.259	0.222	0.303	0.043	0.344	630	0.0080
C-119	Triticale	0.070	0.022	0.262	0.240	0.307	0.033	0.329	751	0.0080
C-120	White clover	0.207	0.025	0.178	0.077	0.235	0.053	0.404	586	0.0081
C-121	White clover	0.271	0.025	0.198	0.195	0.235	0.072	0.201	598	0.0091
C-122	White clover	0.163	0.026	0.182	0.062	0.247	0.053	0.449	538	0.0094
C-123	White clover	0.260	0.022	0.264	0.105	0.264	0.095	0.253	532	0.0099
C-124	White clover	0.204	0.022	0.227	0.130	0.299	0.097	0.246	520	0.0099
C-125	White clover	0.224	0.023	0.215	0.164	0.274	0.086	0.228	500	0.0103
C-126	White clover	0.252	0.027	0.208	0.030	0.244	0.081	0.367	592	0.0108
C-127	White clover	0.285	0.030	0.183	0.140	0.198	0.074	0.275	506	0.0110
C-128	White clover	0.259	0.022	0.198	0.133	0.255	0.075	0.256	542	0.0112
C-129	White clover	0.223	0.042	0.195	0.183	0.260	0.052	0.241	578	0.0112
C-130	White clover	0.258	0.020	0.190	0.037	0.261	0.083	0.340	560	0.0114
C-131	White clover	0.320	0.021	0.257	0.148	0.267	0.083	0.162	538	0.0131

Table SM-2: Chemical composition in kg/kg_{VS}, biogas yield in L/kg_{VS}, and hydrolysis rate constant in h^{-1} of dataset's energy crops for model validation.

Nr.	Common Name	XP	XL	XF	HC	CL	ADL	NFC	Y_B	$k_{h\_0.5}$
V-1	Cup plant	0.091	0.008	0.379	0.081	0.457	0.068	0.295	481	0.0076
V-2	Cup plant	0.072	0.009	0.409	0.105	0.450	0.076	0.289	475	0.0080
V-3	Cup plant	0.050	0.016	0.386	0.064	0.430	0.075	0.365	478	0.0084
V-4	Cup plant	0.076	0.009	0.424	0.153	0.477	0.077	0.209	458	0.0064
V-5	Cup plant	0.074	0.012	0.412	0.135	0.468	0.074	0.237	469	0.0068
V-6	Cup plant	0.101	0.011	0.359	0.070	0.450	0.072	0.296	486	0.0091
V-7	Cup plant	0.070	0.009	0.347	0.078	0.431	0.063	0.349	545	0.0092
V-8	Cup plant	0.055	0.012	0.424	0.130	0.461	0.077	0.264	483	0.0095
V-9	Maize	0.067	0.009	0.270	0.268	0.315	0.033	0.309	720	0.0069
V-10	Maize	0.068	0.018	0.234	0.270	0.268	0.037	0.339	691	0.0078
V-11	Maize	0.063	0.017	0.266	0.258	0.337	0.045	0.280	700	0.0065
V-12	Maize	0.076	0.018	0.229	0.279	0.268	0.030	0.330	698	0.0099
V-13	Maize	0.069	0.024	0.234	0.228	0.290	0.035	0.354	615	0.0092
V-14	Maize	0.061	0.025	0.212	0.199	0.256	0.031	0.429	680	0.0102
V-15	Perennial rye	0.091	0.009	0.376	0.281	0.407	0.075	0.137	561	0.0048
V-16	Perennial rye	0.062	0.011	0.449	0.268	0.460	0.086	0.113	589	0.0053
V-17	Perennial rye	0.047	0.008	0.426	0.300	0.446	0.077	0.122	583	0.0049
V-18	Perennial rye	0.058	0.013	0.429	0.280	0.452	0.091	0.106	574	0.0051
V-19	Perennial rye	0.100	0.013	0.399	0.282	0.410	0.067	0.128	595	0.0066
V-20	Perennial rye	0.107	0.013	0.387	0.287	0.401	0.062	0.130	569	0.0057
V-21	Perennial rye	0.094	0.012	0.309	0.273	0.326	0.052	0.244	602	0.0061
V-22	Perennial rye	0.079	0.011	0.326	0.274	0.351	0.062	0.223	557	0.0062
V-23	Perennial rye	0.088	0.012	0.377	0.310	0.377	0.056	0.157	598	0.0058

V-24	Perennial rye	0.066	0.011	0.415	0.280	0.437	0.061	0.145	637	0.0053
V-25	Perennial rye	0.106	0.012	0.393	0.282	0.418	0.059	0.124	622	0.0059
V-26	Perennial rye	0.089	0.010	0.368	0.270	0.390	0.060	0.181	582	0.0065
V-27	Perennial rye	0.076	0.016	0.372	0.117	0.430	0.075	0.287	644	0.0056
V-28	Perennial rye	0.105	0.012	0.395	0.296	0.417	0.065	0.105	643	0.0059
V-29	Reed canary grass	0.143	0.013	0.309	0.335	0.332	0.058	0.119	525	0.0062
V-30	Rye	0.064	0.014	0.331	0.198	0.352	0.059	0.314	621	0.0073
V-31	Rye	0.048	0.013	0.266	0.185	0.278	0.037	0.438	686	0.0084
V-32	Rye	0.053	0.012	0.221	0.251	0.300	0.051	0.332	627	0.0069
V-33	Rye	0.051	0.012	0.318	0.220	0.300	0.049	0.369	623	0.0088
V-34	Rye	0.065	0.014	0.276	0.220	0.292	0.054	0.355	621	0.0091
V-35	Sorghum	0.081	0.008	0.276	0.314	0.333	0.046	0.218	662	0.0067
V-36	Switchgrass	0.066	0.009	0.333	0.390	0.328	0.061	0.147	544	0.0058
V-37	Switchgrass	0.077	0.014	0.371	0.323	0.394	0.059	0.133	551	0.0048
V-38	Switchgrass	0.081	0.008	0.304	0.375	0.313	0.054	0.168	533	0.0059
V-39	Switchgrass	0.052	0.008	0.366	0.329	0.366	0.065	0.180	545	0.0047
V-40	Switchgrass	0.051	0.007	0.378	0.333	0.370	0.070	0.169	550	0.0048
V-41	Switchgrass	0.081	0.010	0.360	0.357	0.374	0.057	0.121	607	0.0049
V-42	Switchgrass	0.095	0.012	0.253	0.327	0.273	0.036	0.257	662	0.0064
V-43	Switchgrass	0.097	0.012	0.270	0.309	0.293	0.045	0.244	602	0.0066
V-44	Switchgrass	0.136	0.013	0.230	0.344	0.253	0.034	0.219	611	0.0073
V-45	Switchgrass	0.115	0.011	0.250	0.345	0.274	0.038	0.219	593	0.0075
V-46	Switchgrass	0.084	0.012	0.303	0.340	0.313	0.060	0.191	590	0.0046
V-47	Switchgrass	0.091	0.012	0.267	0.328	0.285	0.044	0.241	567	0.0066
V-48	Switchgrass	0.123	0.010	0.264	0.359	0.290	0.038	0.182	541	0.0073
V-49	Virginia mallow	0.093	0.013	0.394	0.199	0.409	0.076	0.210	515	0.0060
V-50	Virginia mallow	0.064	0.016	0.382	0.123	0.399	0.080	0.318	501	0.0081
V-51	Virginia mallow	0.061	0.012	0.366	0.131	0.405	0.072	0.320	508	0.0078

V-52	Virginia mallow	0.058	0.008	0.454	0.162	0.493	0.076	0.203	449	0.0055
V-53	Virginia mallow	0.047	0.011	0.466	0.124	0.506	0.088	0.225	378	0.0054
V-54	Virginia mallow	0.072	0.012	0.324	0.276	0.351	0.055	0.234	452	0.0057
V-55	Virginia mallow	0.081	0.013	0.340	0.112	0.387	0.058	0.348	499	0.0105
V-56	Virginia mallow	0.203	0.014	0.205	0.090	0.250	0.043	0.400	578	0.0126
V-57	Virginia mallow	0.146	0.016	0.260	0.106	0.294	0.057	0.380	556	0.0128
V-58	Virginia mallow	0.153	0.020	0.216	0.098	0.253	0.053	0.423	591	0.0141
V-59	Virginia mallow	0.168	0.018	0.201	0.103	0.216	0.045	0.450	547	0.0151
V-60	Virginia Wildrye	0.133	0.014	0.211	0.244	0.249	0.034	0.325	673	0.0099
V-61	Virginia Wildrye	0.053	0.008	0.322	0.273	0.336	0.062	0.267	544	0.0062
V-62	Wheatgrass	0.073	0.007	0.437	0.310	0.449	0.058	0.103	589	0.0049
V-63	Wheatgrass	0.083	0.008	0.429	0.286	0.431	0.079	0.114	531	0.0047
V-64	Wheatgrass	0.075	0.009	0.416	0.316	0.434	0.059	0.107	635	0.0058
V-65	Wheatgrass	0.075	0.008	0.412	0.306	0.417	0.077	0.116	581	0.0047
V-66	Wheatgrass	0.036	0.007	0.457	0.276	0.446	0.088	0.147	526	0.0040
V-67	Wheatgrass	0.053	0.007	0.428	0.260	0.436	0.080	0.164	558	0.0044
V-68	Wheatgrass	0.088	0.010	0.385	0.325	0.410	0.059	0.109	629	0.0047
V-69	Wheatgrass	0.064	0.008	0.404	0.317	0.419	0.058	0.134	636	0.0047
V-70	Wheatgrass	0.080	0.009	0.422	0.311	0.437	0.056	0.108	597	0.0049
V-71	Wheatgrass	0.076	0.006	0.420	0.274	0.445	0.065	0.134	523	0.0048
V-72	Wheatgrass	0.052	0.007	0.420	0.288	0.439	0.076	0.138	584	0.0043
V-73	Wheatgrass	0.048	0.007	0.439	0.277	0.454	0.081	0.132	559	0.0047
V-74	Wheatgrass	0.065	0.007	0.401	0.221	0.435	0.076	0.196	573	0.0041
V-75	Wheatgrass	0.124	0.012	0.328	0.270	0.359	0.057	0.178	556	0.0081
V-76	Wheatgrass	0.107	0.013	0.327	0.263	0.348	0.040	0.229	594	0.0083
V-77	Wheatgrass	0.045	0.006	0.440	0.291	0.456	0.075	0.127	624	0.0044
V-78	Wheatgrass	0.106	0.014	0.318	0.270	0.338	0.040	0.231	572	0.0082
V-79	Wheatgrass	0.045	0.006	0.442	0.267	0.477	0.078	0.128	601	0.0044

V-80	Wheatgrass	0.166	0.013	0.300	0.322	0.328	0.045	0.126	579	0.0080
V-81	Wheatgrass	0.113	0.010	0.365	0.317	0.375	0.050	0.136	562	0.0067
V-82	Wheatgrass	0.150	0.014	0.317	0.339	0.339	0.058	0.100	580	0.0069
V-83	Wheatgrass	0.119	0.013	0.332	0.305	0.350	0.061	0.153	571	0.0069
V-84	Wheatgrass	0.197	0.012	0.302	0.340	0.322	0.043	0.085	598	0.0079
V-85	Wheatgrass	0.167	0.016	0.280	0.297	0.305	0.046	0.169	598	0.0079
V-86	Wheatgrass	0.183	0.012	0.285	0.296	0.315	0.045	0.148	601	0.0076
V-87	Wheatgrass	0.160	0.015	0.269	0.332	0.285	0.036	0.172	623	0.0081
V-88	Wheatgrass	0.125	0.014	0.308	0.312	0.326	0.050	0.173	485	0.0075
V-89	Wheatgrass	0.124	0.011	0.339	0.327	0.355	0.046	0.138	619	0.0067
V-90	Wheatgrass	0.143	0.013	0.297	0.280	0.316	0.043	0.205	581	0.0077
V-91	Wheatgrass	0.184	0.014	0.282	0.318	0.293	0.040	0.150	602	0.0085

Paper IV

Bioresource Technology 265 (2018) 372–379

Contents lists available at ScienceDirect

Bioresource Technology

journal homepage: www.elsevier.com/locate/biortech

Predicting methane yield by linear regression models: A validation study for grassland biomass

Vasilis Dandikas[a], Hauke Heuwinkel[b], Fabian Lichti[a], Jörg E. Drewes[c], Konrad Koch[c,*]

[a] Institute for Agricultural Engineering and Animal Husbandry, Bavarian State Research Center for Agriculture, Am Staudengarten 3, 85354 Freising, Germany
[b] Department of Agriculture and Food Economy, Hochschule Weihenstephan-Triesdorf, Am Staudengarten 1, 85354 Freising, Germany
[c] Chair of Urban Water Systems Engineering, Technical University of Munich, Am Coulombwall 3, 85748 Garching, Germany

ARTICLE INFO

Keywords:
Biomethane potential
Grassland species
Lignocellulose
Regression model

ABSTRACT

The objectives of this study were to assess and validate previously published prediction models with an independent dataset and to expose the power and limitations of linear regression models for predicting biomethane potential. Two datasets were used for the validation, one with all individual samples and one with the average values of each cultivar. The results revealed similar performances of all four models for the individual samples. The methane yields of the cultivars were predicted more accurately than the methane yields of the individual samples. The grassland specific model predicted the variation in the dataset with an R^2 of 0.84 and the slope of the regression line was equal to 1.0. Linear regression models are suitable to depict the variation in methane yield and for substrate ranking. However, the prediction error of the absolute values may be high since systematic external effects cannot be determined by a regression model.

1. Introduction

Sustainable energy policy and low emission energy production systems are needed in order to reduce the anthropogenic contribution to greenhouse gas emissions and global warming. Moreover, the European Union (EU-28) has to increase the share of renewable energy resources in order to achieve the targets committed in the Paris climate agreement (Liobikienė and Butkus, 2017). Anaerobic digestion for the production and utilization of biogas is considered a cornerstone of both energy transition and circular economy. The biogas sector can provide sustainable electrical energy with a low or neutral carbon footprint. Methane production from lignocellulosic agricultural biomass is considered as an environmentally sound and sustainable form of energy generation (Chandra et al., 2012). However, due to the increasing land use for energy crops production, first generation biomass has been critically discussed. For instance, the increasing maize cultivation in Germany in the recent past has compromised public acceptance of agricultural biogas plants (Kortsch et al., 2015). Recent studies have shown that sustainable alternatives to the use of maize as an energy crop are available in the EU-28 member states, the use of which would assure a continuous progressive development of the European biogas sector (Meyer et al., 2017). For that purpose, second generation biomass is the key parameter for sustainable growth. Roadside grass and agricultural and non-agricultural grasslands represent a high potential of biomass, and its utilization is desirable.

Biogas can be utilized according to the current energy demand, in order to cover electricity production gaps (Mauky et al., 2016). To define an optimal flexible operational concept of a biogas plant, many factors should be taken into account; one of them is the biogas potential of the feedstock. This is accomplished through biomethane potential (BMP) tests; however, the execution of those tests is a time-consuming and costly process. Alternatively, fodder analysis is a standard method to assess the nutritive value of feedstock, with low-cost and reliable analytical methods already available. Thus, models allowing the prediction of biogas and methane yield based only on the chemical composition of a sample are desirable. In this study, four BMP prediction models were selected in order to demonstrate the differences of global and specific models and the influence of the selected regressors on the prediction. Firstly, the model developed by Triolo et al. (2011) was selected. This methane yield (Y_M) prediction model was developed for energy crops, and it was calibrated with ten values. During the calibration of the model, a R^2 of 0.77 was recorded with a relative average estimation error of 6%. Cellulose (CL) and lignin (ADL) were used as regressors. The model depicts only the negative effect of the fibers on methane production. Thus, the model is not suitable for samples with low or no detectable fibers or for samples with extremely high fiber concentrations, since it will overestimate or underestimate the methane production, respectively. The second selected model was developed by

* Corresponding author.
E-mail address: k.koch@tum.de (K. Koch).

https://doi.org/10.1016/j.biortech.2018.06.030
Received 14 April 2018; Received in revised form 9 June 2018; Accepted 11 June 2018
Available online 12 June 2018

Thomsen et al. (2014). There are two versions of this model based on different biomass analysis methods. In this study, the model based on fodder analysis was used. The authors used a dataset consisting of their own data, as well as values from literature. In total 64 samples were used in order to develop a global model for BMP prediction. The relative average estimation error was 18% with a R^2 of 0.97 during calibration (Thomsen et al., 2014). The model was presented as $Y_M = 347$ (CL + HC + R) − 438 ADL, but since residuals (R) were defined as R = 1 − CL − HC − ADL, the regressors of the model can be limited only to ADL. The third selected model was developed by Dandikas et al. (2014), which was also based on two regressors, namely hemicellulose (HC) and ADL. HC was used to express the positive and ADL the negative effect of fibers on biogas production. To develop the model 31 samples were used, and during the calibration a R^2 of 0.83 was recorded with a relative average estimation error of 8%. Finally, a grassland species specific model developed by Dandikas et al. (2015) was utilized. In contrast to the global model of Dandikas et al. (2014), it was proven that for grassland species, crude protein (XP) was also needed to depict the differentiation of the samples on Y_M. The calibration of the model based on 61 samples and the relative average estimation error was 5% with a R^2 of 0.70.

The objective of this study is to assess and validate the previously published grassland species specific model (Dandikas et al., 2015) by an external independent dataset. Moreover, three global prediction models for energy crops were chosen in order to expose the possibilities and limitations of each modeling approach. All four biomethane potential (BMP) prediction models were validated two times by an independent dataset of grassland samples, once with all individual samples and once with the average value of each cultivar.

2. Material and methods

2.1. Substrate

Field experiments have been performed under defined conditions in order to create a dataset with grassland plant species at different developmental stages. The plants were harvested based on the BBCH-scale. The BBCH-scale is a standard two-digit code, to describe the phenological growth stage of plant species (Hess et al., 1997). Six typical agricultural grassland plant species were selected: four grass species (*Lolium perenne*, *Dactylis glomerata*, *Poa pratensis*, *Festuca pratensis*) and two legume species (*Trifolium pratense*, *Trifolium repens*), grown in field plots of 10 m^2. Moreover, four cultivars from the species *Lolium perenne* were tested (Table 1). The crop trial was performed in 2014 in Pulling, Germany. In order to gain information about the chemical composition of the crops along with their development, each cultivar was harvested several times over a certain period in 2014 at distinct phenological stages. The collection of the samples took place at first, second and third growth cycles. During the first growth cycle, five harvest dates were scheduled at defined phenological development stages based on the BBCH-scale, and during the second and third growth cycle, three harvest dates. However, due to weather conditions in 2014, the harvest dates were reduced. In total, 99 samples were scheduled to be harvested; however, only 55 samples could finally be used in the dataset (Table 2). Directly after harvest, the samples were dried in a forced air oven at 40 °C. After the drying process, the samples were ground with a cutting mill to pass a sieve of 10 mm (Retsch SM 200, Haan, Germany) and stored at room temperature.

The validation of the models was performed in two different approaches. The first approach was the validation of the models with each individual sample (N = 55) and the second approach with each cultivar (N = 9). The mean value of seven individual samples was calculated for the cultivars Respect, Sirius and Husar. The mean value of six samples was calculated for the cultivars Sponsor, Preval, Titus and Lirepa. For the cultivars Arvicola and Lato, the mean value of five samples was used. During the first approach, the effect of harvest date (i.e.

Table 1
Harvest dates of the investigated grassland species in 2014; the developmental stage given by BBCH-code in parenthesis.

Species	Cultivar	1. Growth cycle					2. Growth cycle			3. Growth cycle		
		1. HD	2. HD	3. HD	4. HD	5. HD	1. HD	2. HD	3. HD	1. HD	2. HD	3. HD
Lolium perenne	Arvicola	–	–	06. May (51/53)	14. May (55/59)	22. May (65)	–	–	25. Jul. (43/65)	09. Oct. (32/33)	–	–
	Respect	08. May (39)	–	14. May (52)	–	02. Jun. (60)	06. Jun. (42)	17. Jun. (57/59)	24. Jun. (65)	09. Oct. (32/33)	–	–
	Sponsor	14. May (37)	02. Jun. (52)	–	06. Jun. (55)	18. Jun. (59/60)	–	–	17. Jul. (61/65)	09. Oct. (32/33)	–	–
	Sirius	14. May (37/38)	22. May (40)	02. Jun. (51)	06. Jun. (57)	18. Jun. (59/60)	–	–	17. Jul. (61/65)	09. Oct. (32/33)	–	–
Dactylis glomerata	Husar	06. May (39)	14. May (42)	–	22. May (53/55)	06. Jun. (59/60)	23. Jul. (43)	22. Aug. (47)	–	09. Oct. (32/33)	–	–
Poa pratensis	Lato	–	–	–	06. May (55)	26. May (61)	25. Jul. (39)	02. Sep. (47)	–	09. Oct. (32/33)	–	–
Festuca pratensis	Preval	08. May (38)	–	14. May (51)	22. May (57)	26. May (59/60)	23. Jul. (43)	–	–	09. Oct. (32/33)	–	–
Trifolium pratense	Titus	03. Jun. (51)	–	17. Jun. (57)	24. Jun. (65/69)	–	04. Jul. (55)	18. Jul. (60/61)	23. Jun. (63)	–	–	–
Trifolium repens	Lirepa	26. May (51)	02. Jun. (55)	06. Jun. (59)	17. Jun. (65/69)	–	–	–	24. Jun. (69)	–	–	17. Jul. (61/65)

HD: Harvest date.

Table 2
Chemical composition, biogas and methane yield of the 55 samples. The chemical compounds are expressed in g/kg_{VS}, the biogas and methane yield in L/kg_{VS}.

Sample	XP	XF	NfE	XL	ST	RS	HC	CL	ADL	OR	NFC	Y_B	Y_M
G-1	138	189	643	30	ND	263	213	235	26	95	358	676	356
G-2	129	179	660	31	ND	292	198	216	14	120	412	645	345
G-3	108	205	657	30	ND	275	226	247	21	93	368	637	333
G-4	125	236	602	37	ND	146	277	259	62	94	240	547	291
G-5	144	213	629	14	ND	215	272	246	22	86	301	572	308
G-6	152	192	617	39	ND	198	197	261	36	118	315	655	354
G-7	170	178	611	40	ND	198	194	242	36	120	318	708	385
G-8	121	239	617	23	ND	162	191	296	36	171	333	698	374
G-9	220	245	499	36	ND	63	319	286	18	53	121	584	316
G-10	191	251	523	35	ND	57	299	296	26	92	153	577	313
G-11	163	231	571	34	ND	109	291	271	25	107	216	597	322
G-12	152	242	592	14	ND	160	295	267	28	83	243	606	328
G-13	161	185	619	35	ND	217	207	242	25	113	330	629	346
G-14	127	226	623	24	ND	162	166	288	37	192	357	652	349
G-15	149	243	580	28	ND	130	296	289	18	90	220	642	345
G-16	97	265	605	33	ND	108	310	315	25	112	220	594	306
G-17	175	240	543	41	ND	76	300	267	39	90	178	547	298
G-18	158	247	582	13	ND	135	343	270	25	57	192	549	302
G-19	169	176	618	37	ND	225	194	231	22	122	346	663	357
G-20	123	201	649	27	ND	252	206	251	17	125	376	651	342
G-21	130	179	659	32	ND	263	206	228	14	128	391	663	353
G-22	90	214	671	24	ND	261	243	265	16	101	361	662	352
G-23	79	267	630	23	ND	164	282	310	28	114	278	599	319
G-24	192	243	525	40	ND	89	289	274	37	79	168	612	362
G-25	155	204	626	15	ND	224	271	238	22	76	300	620	335
G-26	183	196	577	43	ND	182	215	240	37	101	282	679	377
G-27	164	221	574	41	ND	165	246	253	26	105	270	620	338
G-28	159	193	604	44	ND	187	219	245	36	109	297	696	379
G-29	124	222	619	35	ND	173	219	274	38	137	310	614	330
G-30	108	301	555	36	ND	87	305	321	37	107	194	570	308
G-31	109	327	527	37	ND	24	304	352	39	130	158	536	290
G-32	167	259	557	18	ND	86	304	306	36	83	169	524	288
G-33	182	208	576	33	ND	154	223	258	31	119	273	636	345
G-34	129	248	596	27	ND	141	244	294	39	126	267	603	329
G-35	108	274	594	24	ND	136	322	303	43	64	200	588	319
G-36	120	319	548	13	ND	51	401	358	39	16	68	488	269
G-37	152	264	572	12	ND	123	361	297	28	27	150	572	310
G-38	139	197	631	33	ND	224	222	251	24	108	332	621	336
G-39	134	225	612	29	ND	179	236	274	20	128	307	652	352
G-40	122	227	622	30	ND	172	254	279	21	121	294	662	359
G-41	109	243	621	28	ND	163	254	294	21	132	295	655	351
G-42	131	273	558	38	ND	105	295	282	50	99	204	574	305
G-43	157	251	579	13	ND	122	305	298	22	84	206	545	297
L-1	162	185	631	21	74	116	64	245	51	266	456	577	318
L-2	157	198	627	19	70	109	76	256	55	258	437	540	297
L-3	174	195	609	22	88	99	88	228	52	248	435	524	287
L-4	209	196	570	25	88	78	147	225	59	169	336	539	292
L-5	178	219	581	21	70	99	137	258	51	184	354	537	292
L-6	178	241	557	24	73	79	129	293	64	161	313	535	290
L-7	227	154	588	31	53	95	104	212	51	227	375	639	352
L-8	191	182	602	25	52	132	99	215	41	246	429	585	319
L-9	207	178	590	25	71	98	77	235	53	235	404	586	325
L-10	163	182	629	26	71	90	62	247	53	289	449	538	289
L-11	223	195	540	42	99	55	183	260	52	88	241	578	286
L-12	285	183	503	30	79	60	140	198	74	135	275	506	277

ND: not detectable, G: grass, L: legume.

developmental stage under different weather conditions) was emphasized, while the second approach focused on the differences of the cultivars with respect to the different plant groups (i.e. species).

The field experiment and the plant species selection was organized and performed by the Institute for Crop Science and Plant Breeding of the Bavarian State Research Center for Agriculture in Freising, Germany.

2.2. Fodder analysis

Total solids (TS), crude ash (XA) and volatile solids (VS) were determined gravimetrically according to Standard Methods for the Examination of Water and Wastewater (APHA, 2017). Crude fiber (XF) content was determined by the Weender analysis; moreover, the important fiber fractions (NDF, ADF, ADL) were determined by the Van Soest method (Van Soest and Wine, 1967). Hemicellulose (HC) and cellulose (CL) contents were calculated (HC = NDF − ADF, CL = ADF − ADL). Nitrogen content was determined by the Dumas method and crude protein (XP) was calculated based on the nitrogen content multiplied by 6.25. Crude lipid (XL) content was determined by the extraction method. Starch (ST) content was determined polarimetrically and reducing sugar (RS) content volumetrically. Non-fiber carbohydrates (NFC), nitrogen-free extract (NfE) and organic residue (OR) were calculated (NFC = 100 − XA − XP − XL − NDF, NfE = 100 − XA − XP − XL − XF,
OR = 100 − XA − XP − XL − ST − RS − NDF). All analytical methods were carried out by the Department of Quality Assurance and Analytics of the Bavarian State Research Center for Agriculture in Freising,

Germany.

2.3. BMP test

The biogas yield (Y_B) and methane yield (Y_M) has been determined under laboratory conditions (BMP test). The experiments were performed based on the German guideline VDI 4630 (2016). The inoculum used for the experiments was the effluent of a pilot-scale agricultural biogas plant. The biogas plant was fed with 80% cattle manure and 20% of a dairy cattle feeding mixture (mostly maize and grass silage) at an organic loading rate of 3.0 $kg_{VS}/(m^3 * d)$, with a hydraulic retention time of 19 days. The digester was operated at 38 ± 1 °C. More information about the inoculum used can be found in Dandikas et al. (2015).

The BMP test was performed with a substrate to inoculum ratio of 0.5 ± 0.1 based on VS at 38 ± 0.5 °C. The volume of biogas was measured by Milligascounters (Ritter Apparatebau GmbH, Bochum, Germany). The digester had a working volume of 1.5 L and each sample was tested in triplicate. One gasbag was attached to three digesters (three replicates) and a gas analysis was performed for every 1.5 L biogas produced. The gas analysis was carried out by infrared sensor for CH_4 und CO_2 and by an electrochemical sensor for O_2 (Awite Bioenergie GmbH, Langenbach, Germany). Y_B and Y_M are reported as standard liter (dry gas at 273.15 K and 1013.25 mbar) for each kilogram volatile solids added (L/kg_{VS}).

2.4. Statistical analysis

The predicted values were compared with the measured values from the BMP tests. Descriptive statistic, correlation and variation analyses were performed in order to quantify the relationship among the values. To evaluate and compare the prediction models, the parameters correlation coefficient (r), slope of the regression line (m), statistical bias (average value of the residuals), root mean square error (RMSE) and coefficient of variation of the RMSE (CVRMSE) were calculated.

3. Results and discussion

3.1. Measured biogas and methane yield

For the first validation approach, each individual sample was used. According to the BMP tests, the biogas and methane yields ranged between 488 and 708 L/kg_{VS}, and 269 and 385 L/kg_{VS}, respectively, with a coefficient of variation (CV) of 9% (Table 2). This indicates that the dataset represents a wide range of different qualities of grassland species, making it suitable as an independent dataset for validation. Among all cultivars, a significant difference in Y_B and Y_M was detected with a p value of 0.02 and 0.04, respectively. Within the two plant groups (grasses and legumes), the difference of Y_M was not significant ($p > 0.1$). However, the Y_M between the two plant groups (grasses and legumes) was found to be highly significantly different from each other ($p < 0.01$). For the second validation approach, the mean values for each plant cultivar were used. Statistically, a significant difference was found between the groups in the Y_B and Y_M with a p-value of 0.01. Table 3 shows the mean values of the biogas yield and the chemical composition of the nine cultivars.

3.2. Methane yield prediction of each individual sample

The models of Triolo et al. (2011) and Thomsen et al. (2014) were developed only for the prediction of the methane yield of energy crops. Therefore, only the values of the methane yield were presented during the validation.

Fig. 1 shows the XY-plot of the measured values versus the predicted values of methane yield, and Table 4 summarizes the performance of all four models. The global models of Triolo et al. (2011), Thomsen et al. (2014), and Dandikas et al. (2014) predicted the average Y_M value of the dataset with high accuracy. The average difference between the predicted and measured values (bias) was only 7, −6 and 6 L/kg_{VS}, respectively. The model of Triolo et al. (2011) overestimated the maximum value and underestimated the minimum value, which led to a higher Y_M range prediction. This observation can be explained by the fact that the model starts at the maximum Y_M at 447 L/kg_{VS}, which is reduced with increasing content of CL and ADL. Since the extreme values in the dataset are described by different plant groups (grasses and legumes) and their composition was different (high CL and low ADL content for grasses with a mean value of 274 and 29 g/kg_{VS}, respectively, and low CL and high ADL content for legumes with a mean value of 239 and 55 g/kg_{VS}, respectively), the model overestimated the maximum value and underestimated the minimum value. The influence of ADL content on Y_M seems to be higher than CL content, since the coefficient of variation (CV) for ADL was higher than CL, with CV values of 41% and 13%, respectively. Herrmann et al. (2016) also highlighted the strong negative effect of ADL content on Y_M. The slope of the regression line was at 0.71, and this was the best value among the presented models (Fig. 1). Additionally, the best correlation coefficient (r) between the measured and predicted values was recorded for this model with a value of 0.52 (Table 4). This indicates that the model of Triolo et al. (2011) could explain the variation in the dataset more accurately than the other models. Since the values were relatively closely distributed around the best fit line, the relative average estimation error (CVRMSE) was 11%.

The model of Thomsen et al. (2014), which starts at a much lower Y_M value than the model of Triolo et al. (2011) and reduces Y_M with increasing ADL content, underestimated the maximum value and overestimated the minimum value. This led to a lower Y_M range. The slope of the regression line was 0.19 and the variation among the samples could not be reflected. However, since the values gathered close to the best fit line, the CVRMSE was only 8%, which is the best value among the presented models.

The model of Dandikas et al. (2014) slightly underestimated both the maximum value and the minimum value; however, the range of the values could be well predicted (Table 4). The slope of the regression line was 0.47 and the CVRMSE was 11%. The model developed specifically for grassland samples of Dandikas et al. (2015) showed a bias of 32 L/kg_{VS}, which caused an overestimation of almost all samples. The range was quite well reflected (Table 4). The correlation between the measured and predicted values was 0.44, the slope of the regression line was 0.34, and the CVRMSE was 13%.

The selected models explained only 27% or less of the variation in the dataset, although the CVRMSE was between 8 and 13%. The correlation between the measured and the predicted values can be characterized as weak or moderate, since the r value was between 0.39 and 0.52. A similar observation was reported by Rath et al. (2014) when comparing stoichiometric and empirical biogas yield prediction models using a dataset of maize cultivars. A possible explanation for this might be the fact that each value in this dataset was treated as an individual sample, despite likely not being a representative sample of the specific harvest conditions. Due to microclimate conditions or other external effects, the physico-chemical structure of the sample could probably not be considered representative for the particular specific harvest date or even the phenological stage (actual status versus BBCH-code). This inevitably can lead to high uncertainty. A second explanation for the poor correlation between the measured and predicted values could be that there are many factors that can affect the precision and accuracy of the chemical fodder analysis and the biological BMP tests (Mittweg et al., 2012; Raposo et al., 2011). For both analytical methods, errors of up to 10% are common; this causes a variation within the dataset, which is not determined by the actual composition of the sample. These type of random effects on the sample traits may hinder the detection of a clear relationship between the composition of a sample and its Y_M (Rath et al., 2014, 2013).

Table 3
Average values of each compound and biogas and methane yield of the nine cultivars. The compounds are expressed in g/kg_{VS}, the biogas and methane yield in L/kg_{VS}.

Parameter	Grasses							Legumes	
	Arvicola	Respect	Sponsor	Sirius	Husar	Lato	Preval	Titus	Lirepa
XP	129	167	144	134	145	138	132	176	216
XF	204	225	234	212	245	263	236	206	179
NfE	638	576	592	626	573	577	604	596	575
XL	29	32	29	28	36	22	28	22	30
ST	ND	ND	ND	ND	ND	ND	ND	77	71
RS	238	135	138	211	129	121	161	97	88
HC	237	255	270	242	259	310	261	107	111
CL	241	274	279	257	285	302	280	251	228
ADL	29	29	28	22	36	36	26	55	54
OR	98	106	109	106	110	70	112	214	203
NFC	336	243	249	317	240	192	273	388	362
Y_B	615	632	602	639	606	577	618	542	572
Y_M	326	342	324	346	330	314	333	296	308

ND: not detectable.

3.3. Methane yield prediction of average values of each cultivar

The average values of the chemical compounds can be divided in two groups: a) grass samples (7 cultivars) and b) legume samples (2 cultivars). The average XP, OR, NFC and ADL content of grasses were significantly lower ($p < 0.05$) than those of legumes, whereas the average XF, and HC content of the grass samples were significantly higher ($p < 0.05$) than those of legumes (Table 3). The chemical composition of the two plant groups was dissimilar. This chemical structure influenced the rate of degradation, and even more the biogas production.

Fig. 2 shows the measured values versus the predicted values of the methane yield for the nine cultivars, and Table 5 summarizes the performance of all four models. There was found a strong correlation

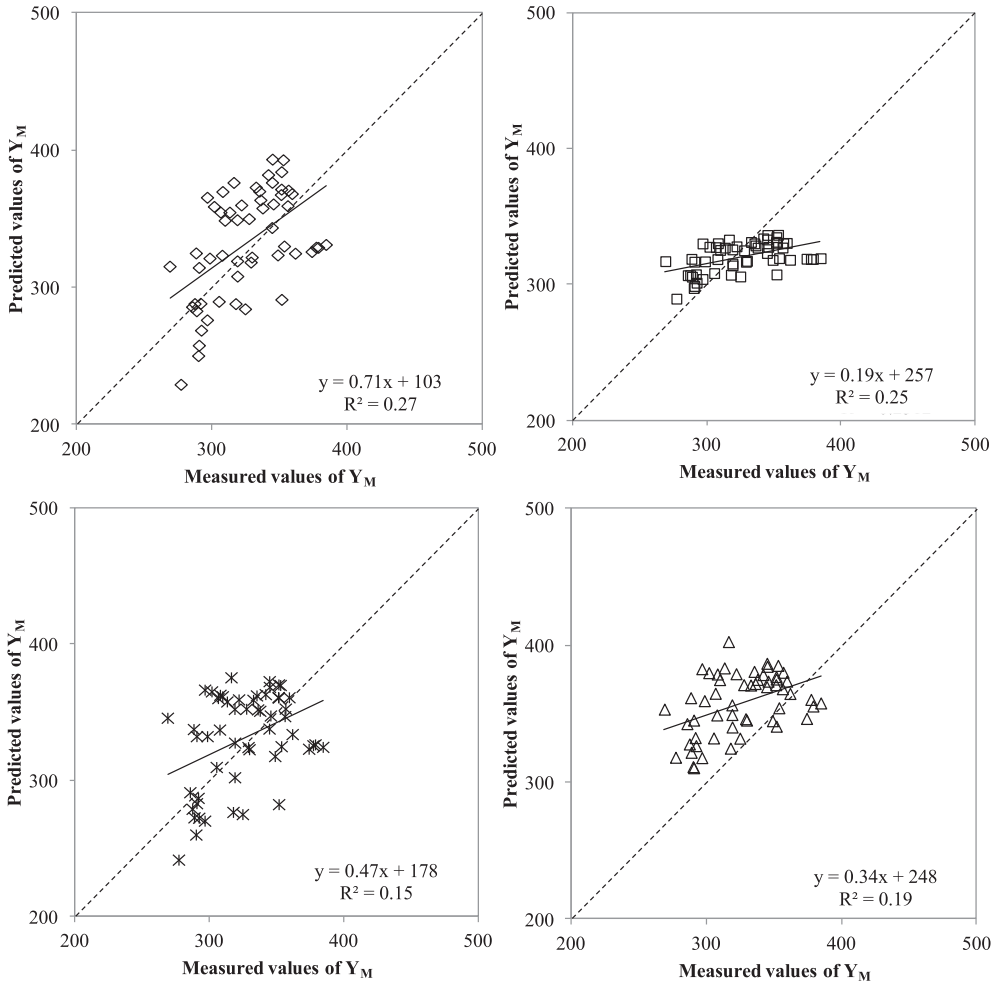

Fig. 1. Measured versus predicted values of methane yield (L/kg_{VS}) from 55 individual samples. Top left: model of Triolo et al. (2011), top right: model of Thomsen et al. (2014), bottom left: model of Dandikas et al. (2014) and bottom right: model of Dandikas et al. (2015).

Table 4
Performance of the tested models to predict the methane yield of the 55 individual samples.

Parameter	Measured values	Triolo et al. (2011)	Thomsen et al. (2014)	Dandikas et al. (2014)	Dandikas et al. (2015)
r	–	0.52	0.50	0.39	0.44
CVRMSE	–	11%	8%	11%	13%
Bias	–	6.76	−5.61	5.68	32.4
Average	325	332	320	331	358
Max	385	393	336	376	403
Min	269	229	289	242	310
Range	116	164	47	134	92

Table 5
Performance of the tested models to predict the methane yield of each cultivar.

Parameter	Measured values	Triolo et al. (2011)	Thomsen et al. (2014)	Dandikas et al. (2014)	Dandikas et al. (2015)
r	–	0.90	0.89	0.85	0.92
CVRMSE	–	6%	3%	6%	10% (2%*)
Bias	–	7.08	−4.84	6.22	32.9
Average	324	331	319	331	357
Max	346	368	330	358	374
Min	296	277	304	274	323
Range	50	91	26	84	51

* After bias correction.

($r > 0.85$) between predicted and measured values for all four models (Table 5), thus more than 70% of the variation of Y_M within the dataset could be explained ($R^2 > 0.7$, Fig. 2). The performance of all four models was markedly improved by using average values; however, differences among the models can be identified. To compare the models and expose their pros and cons, the statistical parameters listed in Table 5 were considered.

The predicted values were significantly correlated with the measured values with r values of 0.90, 0.89, 0.85 and 0.92, for the models of Triolo et al. (2011), Thomsen et al. (2014), Dandikas et al. (2014), and Dandikas et al. (2015), respectively. Again, the global model of Triolo et al. (2011) and Dandikas et al. (2014) performed very similar for the tested dataset. Both models predicted the average value of methane yield close to the measured value and they separated the two plant groups (grasses and legumes) well. Their slopes of the regression lines were again too steep with 1.80 and 1.69, respectively. This reflected the overestimation of the maximum value and the underestimation of the minimum, resulting in an overestimated range of the values. The CVRMSE was 6% for both global models.

The two groups (grasses and legumes) could not be clearly defined in XY-plot by the global model of Thomsen et al. (2014), which showed an overall similar performance as for the individual samples. Although the difference of ADL content between the two plant groups was significant ($p < 0.01$), a model based on ADL alone could not accurately predict the methane yield. The correlation coefficient between the predicted and the measured values was 0.89, but the slope of the

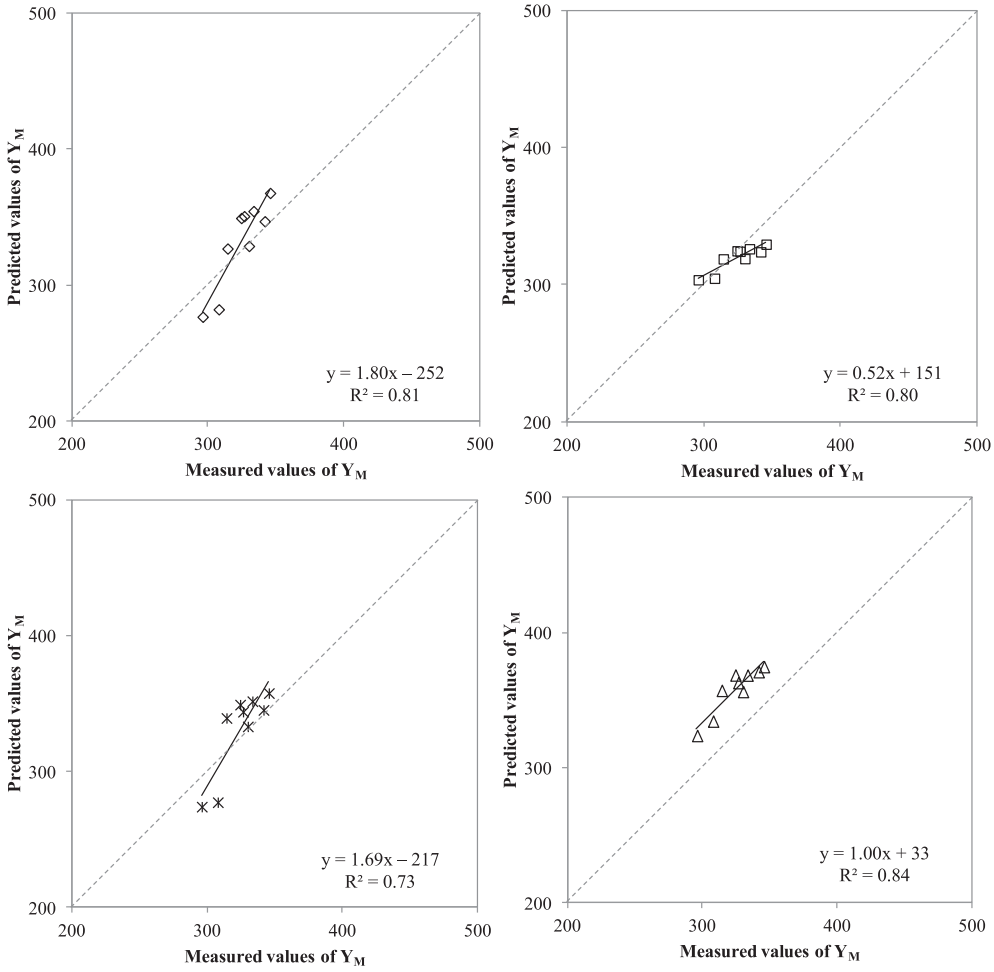

Fig. 2. Measured versus predicted values of methane yield from 9 cultivars (average values). Top left: Model of Triolo et al. (2011), top right: Model of Thomsen et al. (2014), bottom left: Model of Dandikas et al. (2014) and bottom right: Model of Dandikas et al. (2015).

regression line was 0.52. Therefore, the maximum value of the dataset was underestimated and the minimum value was overestimated, resulting in a lower range of the values. However, since the values were close to the best fit line, the CVRMSE was only 3%.

The grassland species specific model of Dandikas et al. (2015) predicted precisely the variation within the dataset, reflected by a slope of the regression line of 1.0 and a correlation coefficient of 0.92. However, a systematic overestimation of each value was observed, since the regression line was parallel to the best fit line and the overestimations were consistently in the same direction. The residuals of the model were recorded between 26 and 44 L/kg_{VS}; the average value of the residuals (bias) was 33 L/kg_{VS}. This systematic deviation from the best fit line caused an CVRMSE of 10%, which is reduced by bias correction to only 2%. It is assumed that the CVRMSE of 2% refers only to the random error of the model. At this point, it should be noted that the bias correction is valid only for this specific dataset. The shift in each datapoint was most likely caused by the different range of the dataset used in the calibration of the model (see Table 2 in Dandikas et al., 2015), which could be explained by differences in the performance of the inoculum and/or the other year and site of harvest. Locher et al. (2005) faced this type of bias during their NIR model validation, as well, which highlights that samples traits may show some systematic year-specific variation.

4. Summary

The single-parameter model of Thomsen et al. (2014) did not fulfill the minimum results expected during validation. This simple concept (only ADL changes Y_M) does not reflect the complex variation of the degradability of the samples, which can be seen by the minor slope of the regression lines. However, this approach can be characterized as a robust mono-causal model.

Using all individual samples for the validation, one key problem was highlighted with all models: the prediction needs reliable data in the sample composition as well as Y_M. The data of all individual samples vary reasonably, causing a low correlation between predicted and measured data. This finding is in line with von Cossel et al. (2018), who also showed that for the prediction of individual samples, r values between 0.12 and 0.51 were recorded. The authors also highlighted the importance of the crop-specific intercepts (bias correction) for high model accuracy. The correlation between predicted and measured data was increased by using average values of each cultivar. In this case, the grassland specific model of Dandikas et al. (2015) proved to be better than the other tested models. However, this positive observation is simultaneously a drawback, since the differentiation between harvest dates (i.e. developmental stages of the plants) was kept hidden within the values. The good prediction with the Dandikas et al. (2015) model means that the model described the plant cultivars better than the changes caused by aging of the plants. A possible explanation is that the aging of these plants was less important for Y_M generation, at least within the observed growth stages, than for the use for fodder production (Tallowin, 1999). Mast et al. (2014) showed that the harvest dates of perennial crops did not always significantly influence the Y_M. Wahid et al. (2015) also reported, that a significant difference on Y_M was not always observed among the harvest dates within a growth cycle of grassland species. In contrast, the Y_M of grassland plants harvested from May to February decreased substantially with a later harvest (Herrmann et al., 2014) Another possible explanation could be that the model may hide the effect of growth stage and this effect was related to the random error of the model. Once again, the results support the general limitation of empirical models that, specific models could differentiate minor changes in the samples only by a distinct calibration. On the other hand, the specific calibration of a model will reduce the robustness and thereby the range of application of the model. For the presented data the models of Dandikas et al. (2014) and Triolo et al. (2011) can be used to rank samples according to Y_M into three levels (low, medium, high) even for the dataset with the values of each individual sample.

5. Conclusions

The linear regression models could not explain more than 27% of the dataset variation of all individual samples; however, they can be used to rank feedstocks. Using the average values of each cultivar the grassland species specific model could predict the variation of the dataset more precisely than the global models. Furthermore, using these average values minimizes the prediction error and improves the estimation of the differences among the samples. External factors may affect the physical–chemical structure of the plants; therefore, a calibration dataset should be updated frequently.

Acknowledgements

This project was funded by the Bavarian State Ministry of Food, Agriculture and Forestry, Germany and the Bavarian State Ministry of Economic Affairs and Media, Energy and Technology, Germany, Project code: BE/15/06.

References

APHA, 2017. Standard Methods for the Examination of Water and Wastewater, twenty third ed. American Public Health Association, Washington, DC.
Chandra, R., Takeuchi, H., Hasegawa, T., 2012. Methane production from lignocellulosic agricultural crop wastes: a review in context to second generation of biofuel production. Renewable Sustainable Energy Rev. 16, 1462–1476.
Dandikas, V., Heuwinkel, H., Lichti, F., Drewes, J.E., Koch, K., 2015. Correlation between biogas yield and chemical composition of grassland plant species. Energy Fuels 29, 7221–7229.
Dandikas, V., Heuwinkel, H., Lichti, F., Drewes, J.E., Koch, K., 2014. Correlation between biogas yield and chemical composition of energy crops. Bioresour. Technol. 174, 316–320.
Herrmann, C., Idler, C., Heiermann, M., 2016. Biogas crops grown in energy crop rotations: linking chemical composition and methane production characteristics. Bioresour. Technol. 206, 23–35.
Herrmann, C., Prochnow, A., Heiermann, M., Idler, C., 2014. Biomass from landscape management of grassland used for biogas production: effects of harvest date and silage additives on feedstock quality and methane yield. Grass Forage Sci. 69, 549–566.
Hess, M., Barralis, G., Bleiholder, H., Buhr, L., Eggers, T., Hack, H., Stauss, R., 1997. Use of the extended BBCH scale—general for the descriptions of the growth stages of mono; and dicotyledonous weed species. Weed Res. 37, 433–441.
Kortsch, T., Hildebrand, J., Schweizer-Ries, P., 2015. Acceptance of biomass plants – results of a longitudinal study in the bioenergy-region Altmark. Renewable Energy 83, 690–697.
Liobikienė, G., Butkus, M., 2017. The European Union possibilities to achieve targets of Europe 2020 and Paris agreement climate policy. Renewable Energy 106, 298–309.
Locher, F., Heuwinkel, H., Gutser, R., Schmidhalter, U., 2005. Development of near infrared reflectance spectroscopy calibrations to estimate legume content of multispecies legume-grass mixtures. Agron. J. 97, 11–17.
Mast, B., Lemmer, A., Oechsner, H., Reinhardt-Hanisch, A., Claupein, W., Graeff-Hönninger, S., 2014. Methane yield potential of novel perennial biogas crops influenced by harvest date. Ind. Crops Prod. 58, 194–203.
Mauky, E., Weinrich, S., Nägele, H.-J., Jacobi, H.F., Liebetrau, J., Nelles, M., 2016. Model predictive control for demand-driven biogas production in full scale. Chem. Eng. Technol. 39, 652–664.
Meyer, A.K.P., Ehimen, E.A., Holm-Nielsen, J.B., 2017. Future European biogas: animal manure, straw and grass potentials for a sustainable European biogas production. Biomass Bioenergy.
Mittweg, G., Oechsner, H., Hahn, V., Lemmer, A., Reinhardt-Hanisch, A., 2012. Repeatability of a laboratory batch method to determine the specific biogas and methane yields. Eng. Life Sci. 12, 270–278.
Raposo, F., Fernández-Cegrí, V., De la Rubia, M.A., Borja, R., Béline, F., Cavinato, C., Demirer, G., Fernández, B., Fernández-Polanco, M., Frigon, J.C., Ganesh, R., Kaparaju, P., Koubova, J., Méndez, R., Menin, G., Peene, A., Scherer, P., Torrijos, M., Uellendahl, H., Wierinck, I., de Wilde, V., 2011. Biochemical methane potential (BMP) of solid organic substrates: evaluation of anaerobic biodegradability using data from an international interlaboratory study. J. Chem. Technol. Biotechnol. 86, 1088–1098.
Rath, J., Heuwinkel, H., Herrmann, A., 2013. Specific biogas yield of maize can be predicted by the interaction of four biochemical constituents. BioEnergy Res. 6, 939–952.
Rath, J., Heuwinkel, H., Taube, F., Herrmann, A., 2014. Predicting specific biogas yield of maize-validation of different model approaches. BioEnergy Res. 1–11.
Tallowin, Jefferson, 1999. Hay production from lowland semi-natural grasslands: a review of implications for ruminant livestock systems. Grass Forage Sci. 54, 99–115.

Thomsen, S.T., Spliid, H., Østergård, H., 2014. Statistical prediction of biomethane potentials based on the composition of lignocellulosic biomass. Bioresour. Technol. 154, 80–86.
Triolo, J.M., Sommer, S.G., Møller, H.B., Weisbjerg, M.R., Jiang, X.Y., 2011. A new algorithm to characterize biodegradability of biomass during anaerobic digestion: influence of lignin concentration on methane production potential. Bioresour. Technol. 102, 9395–9402.
Van Soest, P.J., Wine, R.H., 1967. Use of detergents in the analysis of fibrous feeds. IV. Determination of plant cell-wall constituents. J. Assoc. Off. Agric. Chem. 50, 50–55.
VDI, 2016. Guideline VDI 4630 – Fermentation of organic materials Characterisation of the substrate, sampling, collection of material data, fermentation tests. The Association of German Engineers, Düsseldorf, Germany.
von Cossel, M., Möhring, J., Kiesel, A., Lewandowski, I., 2018. Optimization of specific methane yield prediction models for biogas crops based on lignocellulosic components using non-linear and crop-specific configurations. Ind. Crops Prod. 120, 330–342.
Wahid, R., Ward, A.J., Møller, H.B., Søegaard, K., Eriksen, J., 2015. Biogas potential from forbs and grass-clover mixture with the application of near infrared spectroscopy. Bioresour. Technol. 198, 124–132.